



## Building the European database of cancer registry data

Eva Steliarova-Foucher

International Agency for Research on Cancer  
Lyon, France

**E  
N  
C  
R**

EUROPEAN NETWORK  
OF CANCER REGISTRIES



### History

- Population-based cancer registries in Europe since 1950s
- Organised in a Network since 1989 (EU funds)
- Currently about 200 members
- Recommendations, courses, consultancies
- EUROCIM
  - Software 1999, 2000, 2001
  - database

International Agency for Research on Cancer



**E  
N  
C  
R**

EUROPEAN NETWORK  
OF CANCER REGISTRIES

## European cancer database - rationale

- Need for an authoritative source of European cancer data
  - Standardised
  - Comparable
  - Complete
  - Informative
- To be used for:
  - Benchmarking and monitoring of cancer burden
  - Planning public health resources
  - Evaluation of strategic decision
  - Research
  - Improvement of the cancer management outcome

International Agency for Research on Cancer



## European cancer database – the added value

- Conformity to international standards
- Standardised and comparable output
- Easy access
- Sharing resources
- Combining expertise
- Exponentiating knowledge

International Agency for Research on Cancer



## European cancer database – the barriers

- Multi-lingual Europe
- Differences in organisation of public health system
- Variations in registration procedures
- Variable infrastructures used by the registries
- Use of local 'standards'
- Data confidentiality requirements
- Registries reluctance to provide the data

International Agency for Research on Cancer



## Europe against Cancer: Optimisation of the Use of Registries for Scientific Excellence in research

[www.eurocourse.com](http://www.eurocourse.com)

1 April 2009 – 31 March 2012

ERA-NET Seventh Framework Programme of the DG Research of the EC

A major project of the ENCR

International Agency for Research on Cancer



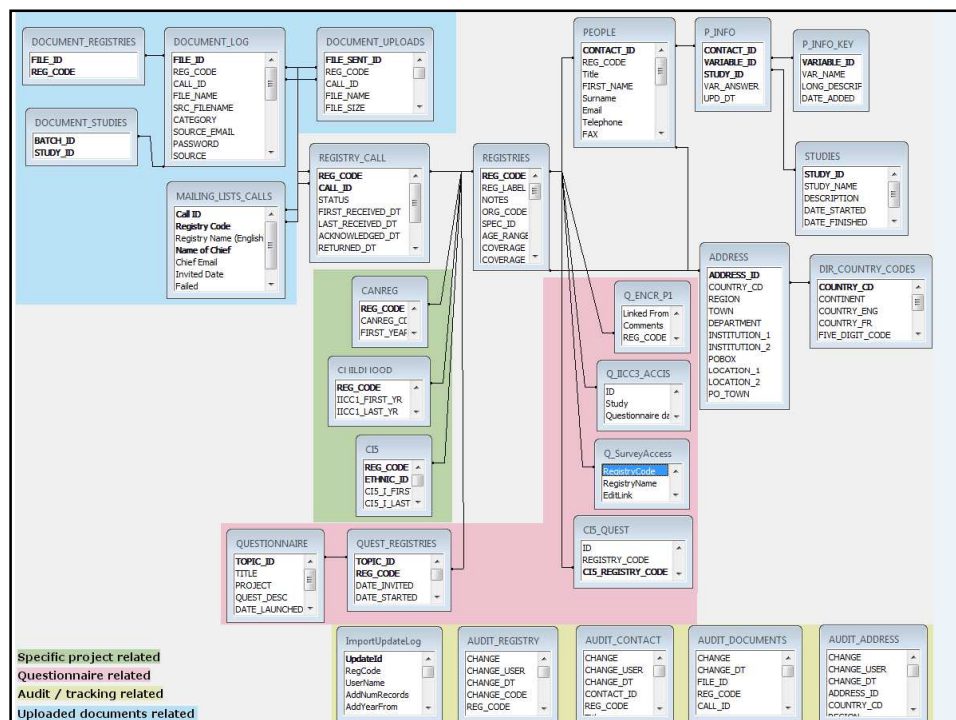
## EUROCOURSE – database-related deliverables

- D 4.1 Public access dataset for routine situation analyses
- D 4.2 Research dataset with access restricted to the registries and researchers
- D 4.5 Development of a European Cancer Observatory: the formation of a comprehensive programme of work on cancer intelligence

## European cancer database – construction

1. Database design
2. Call for data
3. Data submission
4. Central data processing
5. Data validation
6. Consolidation of the common database
7. Dissemination

- Brand new database
- Stored on SQL Server
- Complex structure with multiple relationships:
  - Registries
  - Individual collaborators
  - Studies
  - Data calls
  - Files
  - Surveys
  - Processes



## 1. Database design



- Automation
  - Data submission/reception
  - Process monitoring
  - Archiving
- Efficiency
  - Less staff time
  - Faster processing
- Reduction of error
  - Transcription
  - Process follow-up
- Increased security
  - Password protected access
  - Less human intervention
- Versatility and sustainability
  - Inclusive of on-going and future activities

International Agency for Research on Cancer



## 2. Call for data



- Working group to prepare a call
- 21<sup>st</sup> May 2010: ENCR call for data
  - Invitation by a letter to the cancer registries
  - Invitation by an e-mail
  - Invitation through the ENCR web-site
- On-line electronic submission
  - First use of the portal
  - Successful transfer of all files so far through portal
  - No other modes of submission used
- Deadline: 30<sup>th</sup> June 2010
- Alignment with the EURO CARE call

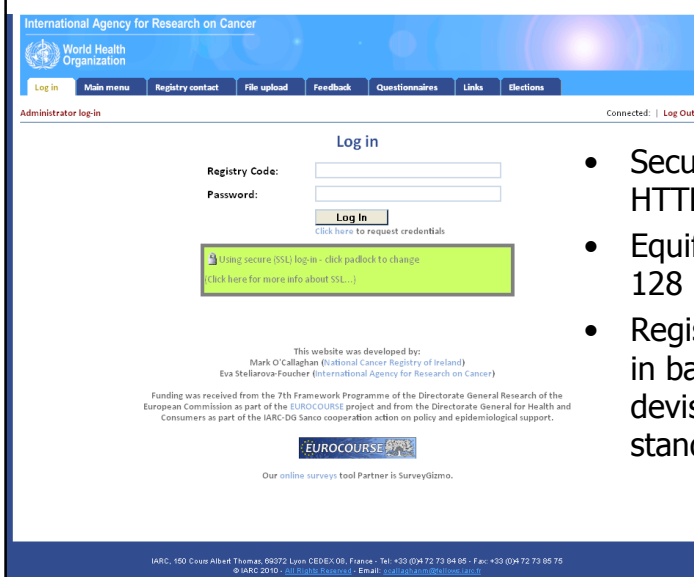
International Agency for Research on Cancer



## 2. Call for data

- Cancer cases file (45 variables)
  - ◆ Person: IDs, sex, age, dates
  - ◆ Tumour: ICD-O-3, stage, treatment
  - ◆ Follow-up information
- Mortality data (sub-national registries)
- Life-table files
- Population data
- Coding schemes

## 3. Data submission - Registries portal



The screenshot shows the login interface of the EUROCOURSE Registries portal. At the top, there is a header with the International Agency for Research on Cancer (IARC) and World Health Organization (WHO) logos, and a navigation menu with links like 'Log in', 'Main menu', 'Registry contact', 'File upload', 'Feedback', 'Questionnaires', 'Links', and 'Elections'. Below the header, there is a 'Log in' section with fields for 'Registry Code:' and 'Password:', and a 'Log In' button. A green box highlights a message: 'Using secure (SSL) log-in - click padlock to change (Click here for more info about SSL...)'. Below this, there is a footer with information about the website's development by Mark O'Callaghan and Eva Stelarcova-Foucher, and funding from the 7th Framework Programme of the Directorate General Research of the European Commission. The EUROCOURSE logo is also present in the footer.

- Secure log-in using HTTPS / SSL
- Equifax high-grade 128 bit encryption
- Registry-specific log-in based on newly devised coding standards

## 4. Central data processing

### Sorting & viewing of ENCR data at IARC

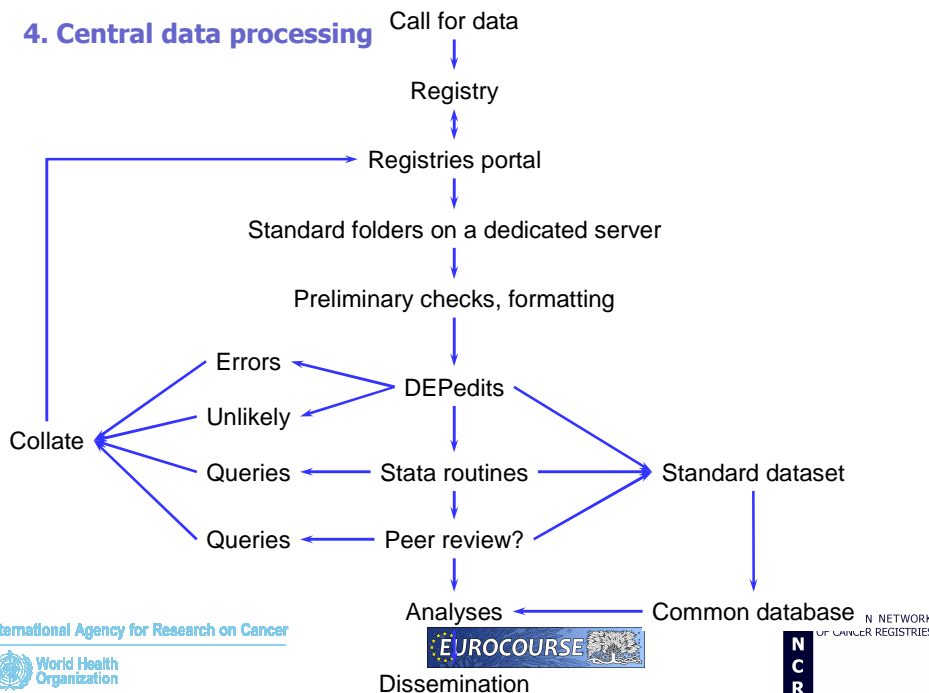
The screenshot shows the 'Records' section of the IARC interface. It includes a 'Select registry code' dropdown set to 'ALL REGISTRIES'. Below this is a table with columns: Registry Code, File name, Call, and Category. The table lists various registries like EUSLUN10, EUSPABCO, and EUSPANAV with their respective file names and categories. On the left, there are filters for 'Call for data' (set to 'ENCR call 2010'), 'Document type', 'Status', and 'Study'. At the bottom, there are fields for 'Link to file', 'Status' (set to 'Acknowledged'), 'Received date', 'Checked date', and 'Finalized date'. A status message at the bottom left says 'You have filtered data for [CALL\_ID] = 1000'.

- Descriptive emails are sent to CIN staff when files are uploaded
  - This occurs at the same time as the registry receives their confirmation
- Files are conveniently named, stored and categorised for easy access
- History and details of files uploaded can be viewed by CIN staff through the new collaborators database (built in Access - VB).
- Reports of the status

RK  
IES

R

## 4. Central data processing





## Status of submissions on 10 June 2011 cancer files

	Received	Expected	Total
Registries submitting	99	57	156
Including 'superregistries'	4	1	5
Including regional registries	46	15	61
Including specialised	2		2
Including childhood	5	6	11
Total registries	95	56	151
Total general registries	87	51	138

International Agency for Research on Cancer



## Status of submissions on 10 June 2011 Cancer records

	N	%
Records submitted	18,413,386	100
Records processed	15,962,054	86.7
Records excluded	53,940	0.3
Records included in summary statistics	15,908,114	86.4

International Agency for Research on Cancer



## Status on 10 June

	Submitted		Processed %	Expected	
	National	Regional		National	Regional
Austria		2	50	1	
Belarus			-	1	
Belgium	1	1	50		
Bulgaria	1		100	-	
Croatia	1		100	-	
Cyprus	1		100	-	
Czech Republic			-	1	
Denmark	1		100	-	
Estonia	1		0	-	
Finland	1		100	-	
France			-		15
Germany		10	50		8
Gibraltar	1		0	-	
Iceland	1		100	-	
Ireland	1		100	-	

International Agency for Research on Cancer



EUROPEAN NETWORK  
OF CANCER REGISTRIES

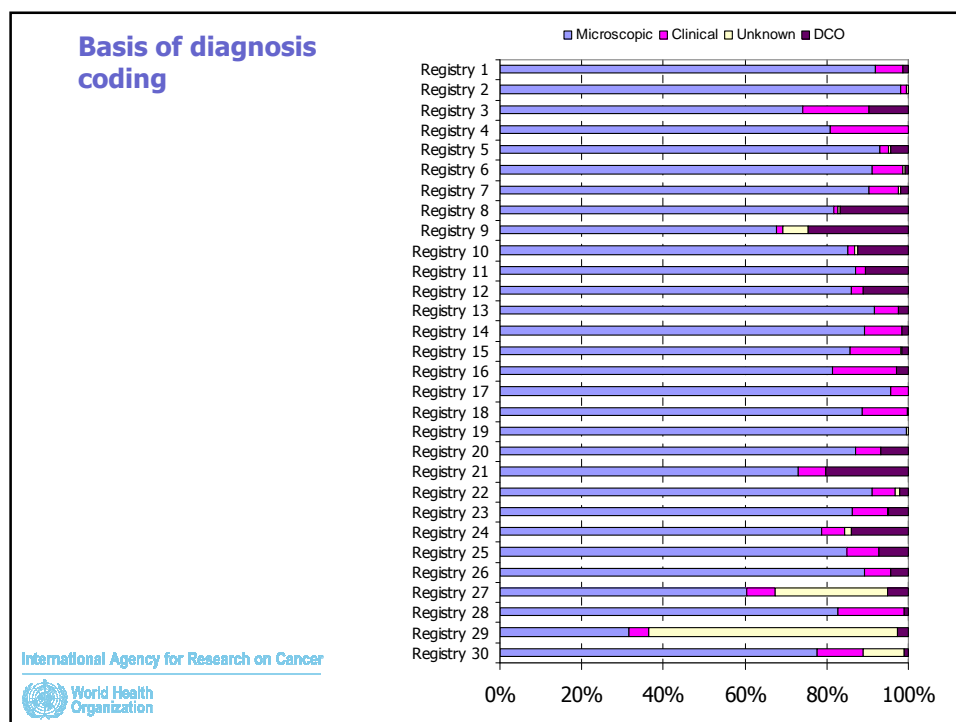
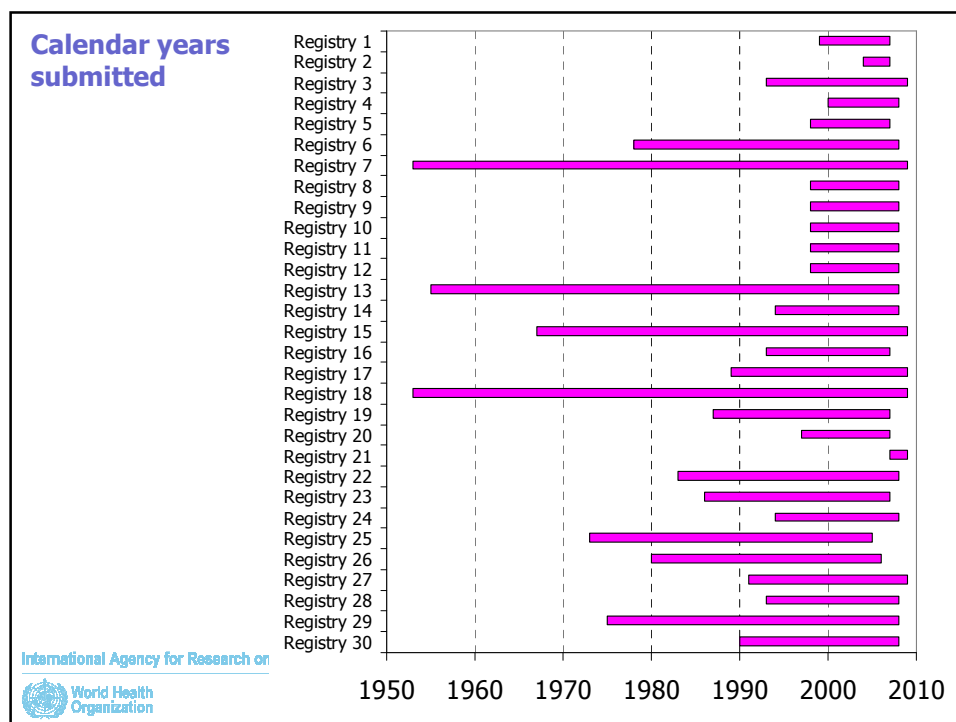
## Status on 10 June ctnd.

	Submitted		Processed %	Expected	
	National	Regional		National	Regional
Italy		28	96		
Lithuania	1		100	-	
Netherlands	1	1	50	-	
Norway	1		100	-	
Poland		1	0		3
Portugal		2	100		2
Romania		2	50		1
Serbia			-		2
Slovakia	1		0	-	
Slovenia	1		100	-	
Spain		6	67		9
Sweden	1	6	0	-	
Switzerland		1	0		9
Turkey		1	0		1
UK	1	11	100	-	

International Agency for Research on Cancer

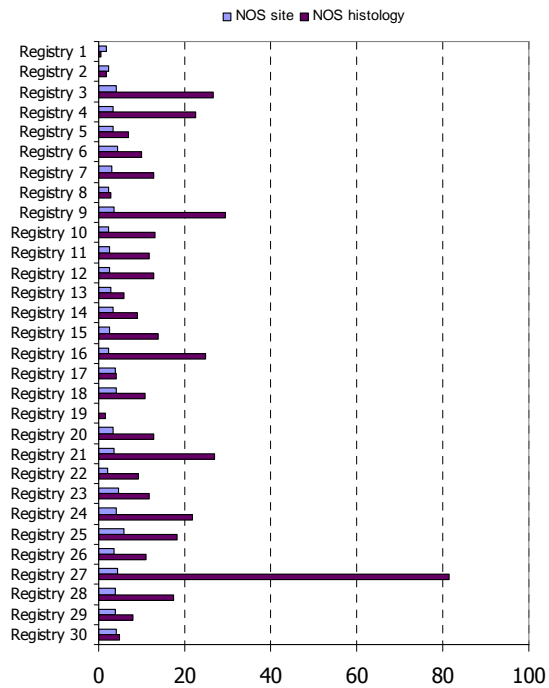


EUROPEAN NETWORK  
OF CANCER REGISTRIES



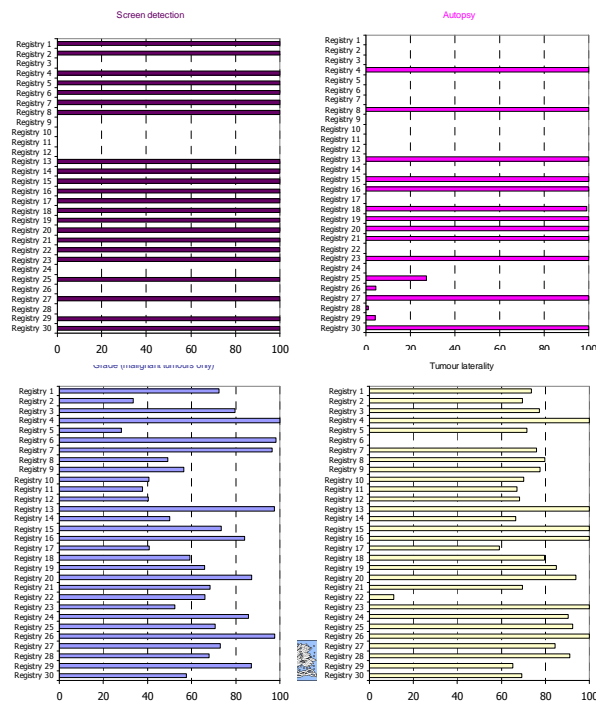
**Unspecified sites  
C26, C39, C48, C76, C80  
Unspecified histologies  
M-8000 – M8005**

International Agency for Research on Cancer

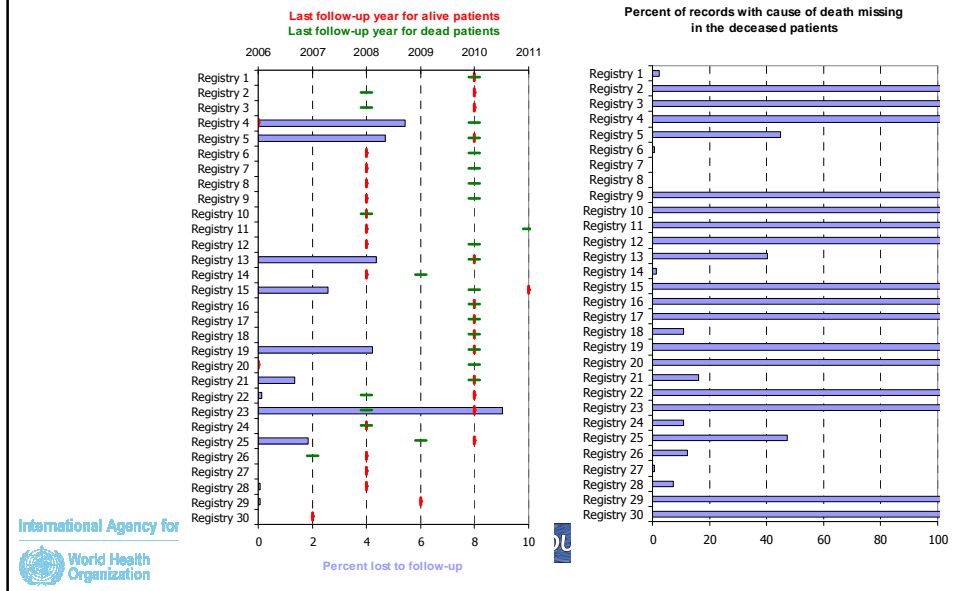


**Percentage  
of records  
with missing  
information**

International Agency for Research on Cancer

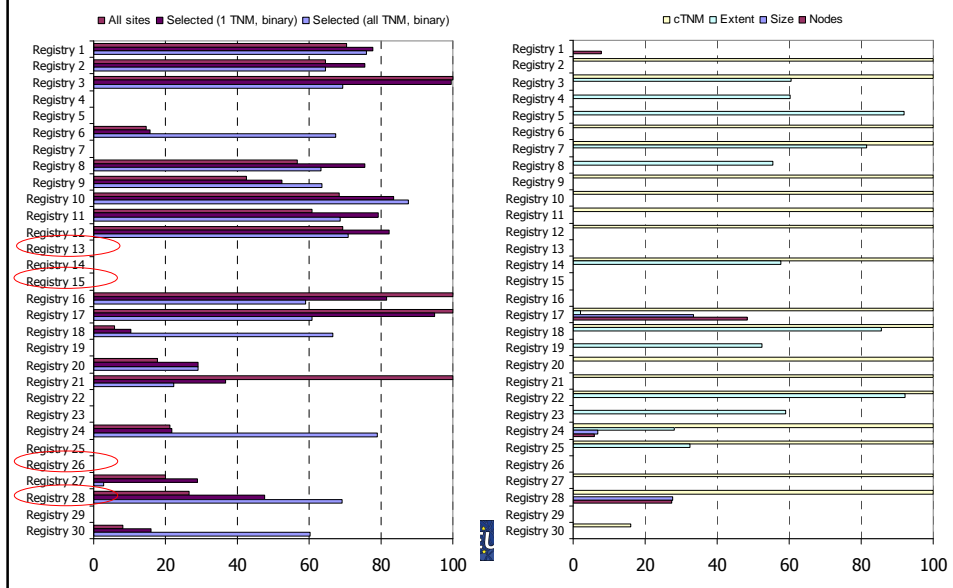


## Follow-up information availability

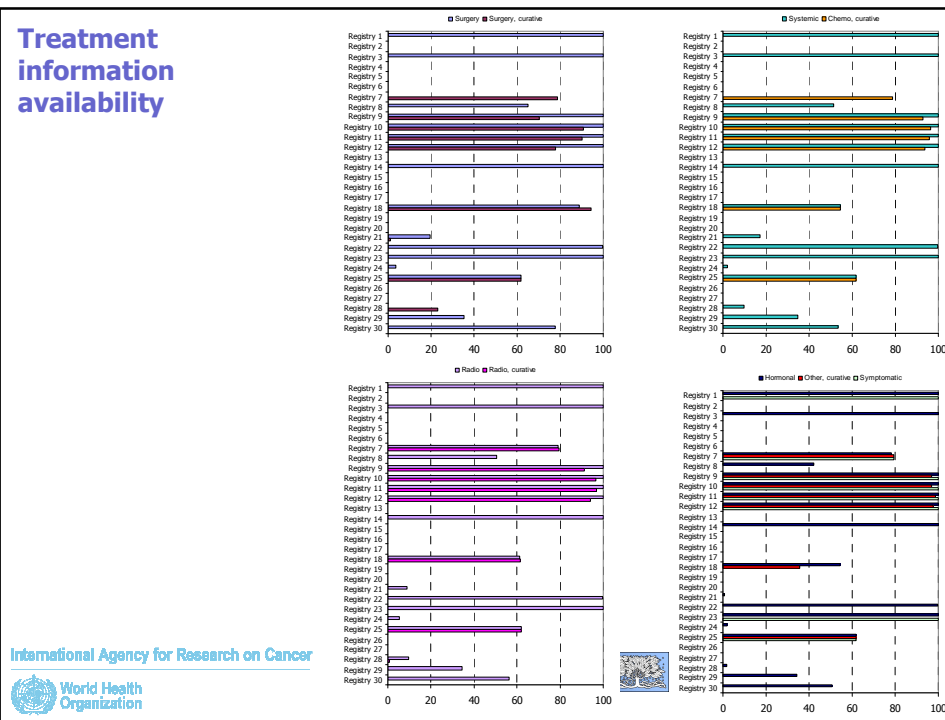


## Stage information availability

Selected sites (ICD=10): C16, C18, C19, C20, C43, C50, C53, C54, C56, C61, C62, C64, C67, C73



## Treatment information availability



## 7. Dissemination – proposal of a WG

- Aggregated data
  - 181 cancer entities by sex, 5-year age groups and the 5-year period as defined as CI5
- Individual records with selected variables
  - ◆ Geographical area
  - ◆ Tumour sequence number
  - ◆ Sex at birth
  - ◆ Month and year of incidence
  - ◆ Age in years
  - ◆ ICD03 Topography
  - ◆ ICD03 Morphology
  - ◆ ICD03 Behaviour
  - ◆ Basis of diagnosis
  - ◆ Grade
  - ◆ Laterality
  - ◆ Vital status at last contact
  - ◆ Month and year of end of follow-up
  - ◆ Duration of survival in days
  - ◆ Cause of death
  - ◆ ICD edition coding the cause of death
- Individual records with all variables

## 7. Dissemination – online search and download

International Agency for Research on Cancer  
World Health Organization  
EURO-CIM  
<http://eurocim.iarc.fr>  
EUROCAN logo

**Search Common Database**  
Presentation of incidence, mortality and survival data in various tables or graphs.  
The tables can be exported to a text file or saved as a PDF file.

**Search Cancer Records**

**Location of case**

☐ Country:   
☐ Registry:

**About the patient**

☐ Ethnic group:

☐ Sex: ☐ Male ☐ Female ☐ Other

☐ Age:  -->

☐ Vital Status: ☐ Alive ☐ Dead ☐ Unknown

☐ Cause of death:

International  
World Health  
Organization

NETWORK  
REGISTRIES

## 7. Dissemination – online search and download

**About the tumour**

☐ Year of incidence:  -->

☐ ICD-O3 Topography:

☐ ICD-O3 Morphology:

☐ ICD-O3 Behaviour: ☐ Benign ☐ Uncertain ☐ In situ ☐ Malignant ☐ Unknown

☐ ICD-O3 Grade: ☐ 1,I ☐ 2,II ☐ 3,III ☐ 4,IV ☐ 5,Tcell ☐ 6,Bcell ☐ 7,0cell ☐ 8,Nkcell ☐ NOS

Home - IARC - 7<sup>th</sup> Framework Programme - EURO-COURSE - ENCR - [ocallaghan@fellows.iarc.fr](mailto:ocallaghan@fellows.iarc.fr)  
© IARC 2011 - All Rights Reserved - Email: [www@iarc.fr](mailto:www@iarc.fr)

Int  
World Health  
Organization

ORK  
RIES

## 7. Dissemination – data sharing Registries survey based on 91 respondents

		Grouped data %	Individual records %
ENCR	controlled mechanism	70.3	71.4
CR	controlled mechanism	63.7	50.5
Anyone	controlled mechanism	41.8	34.1
ENCR	password	48.4	36.3
CR	password	44.0	28.6
Anyone	password	27.5	18.7
Anyone	no limits	27.5	6.6
Other		7.7	16.5

International Agency for Research on Cancer



## Conclusions



- Construction of the European database in progress
- Access to the data under discussion
- Communication with the registries planned
- Output is being developed
- Good compliance by the registries
- Need to work on data comparability
- Review of the Call for data

International Agency for Research on Cancer

