Automated research and mining of real-life hospital cancer care data

A proposal of functional solution realized in project “Educational platform I-COP”

THIS PROJECT IS SUPPORTED BY THE EUROPEAN SOCIAL FUND AND THE STATE BUDGET OF THE CZECH REPUBLIC.

I-COP

MU IBA

Presentation is divided into 3 parts

- I. Briefly about the national project “Educational and informative platform for comprehensive cancer centers - CCCs”
- II. Software application support of the project - draft and development of the application for automated research and mining of real-life hospital cancer care data
- III. Predicted possibility of using the application as an information support for screening programs in hospitals (colorectal carcinoma example)

Part I

Project “Educational and informative platform for comprehensive cancer centers”

Project background

- Realization of the project is supported by the European social fund and the state budget of the Czech Republic. This two-year project was started on 1 February 2012.

- Project is lead by the Masaryk University Brno, Institute of Biostatistics and Analyses in cooperation with 13 partner hospitals

- The professional guarantee of this project is the Czech Society for Oncology

Goals of the project

- To build up an educational and informational support for the oncologists
  - Provide workshops, seminars and conferences
  - Central education portal with on-line education materials and live reports will be available on www.icop.cz
  - **Scientific publications and data analyses will be the substantial outcome of the project**

- Bring/reveal new data sources for the oncologists
  - The administrative and epidemiological data about the cancer care in all the partner hospitals will be processed
  - Thanks to the involvement of most of the CCCs, the project will be able to analyze representative portion of national data about cancer care in the Czech republic.
  - **The information support of the CCCs and the new data source will be implemented as a software data warehouse application providing analytical output over the data of involved CCCs**

Part II

Software application support of the project - draft and development of the application for automated research and mining of real-life hospital cancer care data

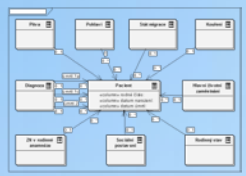
Real-life hospital cancer care data

- Data from Hospital Information System(s)
  - HIS's differ among CCCs
  - Contains a lot of information in non-parametric form
  - Cooperation of the supplier of the HIS, additional costs for the CCCs
  - Only a supportive role

- Use of the more universal data sources
  - **Medical Insurance Records**
  - **National Cancer Registry**

Data integration

Czech Cancer Registry (CCR)



Medical insurance reports

?

Tumor diagnosis

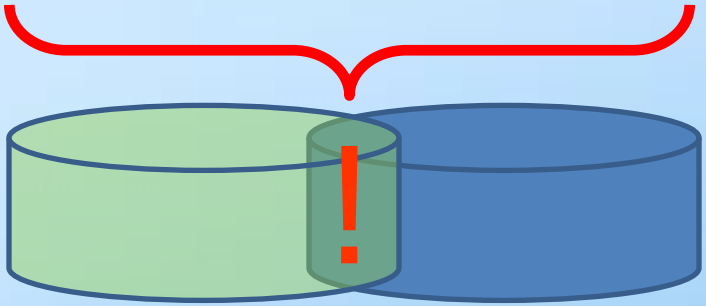
TNM classification

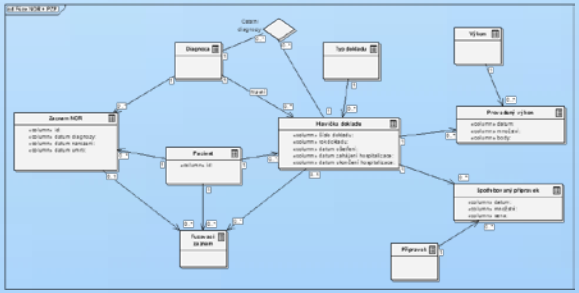
Clinical stage

Patients treatment

Hospital processes and procedures

Approximate costs





A new information-rich resource

=

I-COP DATABASE

I-COP SOFTWARE

Conceptual draft of the proposed application

Presented outcomes are based on realistic, but not a real data in order to protect privacy of patients and partner hospitals.

Proposed I-COP components

I-COP Database – integration of clinical information on primary diagnosis and treatment of oncologic patients from the Czech Cancer Registry (CCR) data and administrative data from medical insurance records will allow us to create information-rich merged database that offers completely new perspectives on the process of the oncologic patients' treatment. This designed I-COP database is a fundament for all I-COP tools described bellow.

I-COP Reporting – I-COP Reporting will provide physicians and other hospital workers with a set of predefined, although still customizable views of the I-COP database. These views will be realized through the secure web-based reporting tool and it will primarily offer the basic analyses of patients structure, treatment load and region-related data of given facility. The outcomes could be directly compared with the national reference values.

I-COP Browser – secure web-based tool that will allow advanced users to define their own, very general views of the raw primary data from I-COP database. The primary part of this tool will be designed as a pivot table with defined selectable dimensions. Each group of patients displayed in this pivot table could be selected by variety of filters.

I-COP analyses – the on demand analyses that cannot be implemented directly in the standard I-COP SW tools. These will be prepared as a standalone output focused on the specific topic of the cancer care.

I-COP REPORTING

I-COP **Cancer Care Information Center**

Home | Reporting | Browser | Analyses | Help | Info | log in

[Introduction] | Epidemiology of cancer | Detection of diseases | Patients of Hxx in NOR | Treatment load of Hxx

I-COP Reporting provides physicians and other hospital workers set of predefined, but customizable views into the database of I-COP. These views are realized through web-based reporting and mainly offer analysis of cancer patients structure, regional related data of given facility and standardized predefined patient information sets compared with average values in the whole country.

Epidemiology of cancer

Descriptive data on the epidemiology of cancer from NOR - incidence, mortality and prevalence in the study population (selected counties or districts).

Detection of diseases

Characteristics of diagnosed patients in the study population - representation of gender, age of the patients, stage detection.

Patients of Hxx in NOR

Characteristics of cancer patients diagnosed, treated or dispensarized in Hxx - regional distribution, representation of diagnosis, stage of disease.

Treatment load of Hxx

Characteristics of cancer patients treated in Hxx (from reported treatment data and NOR) - regional distribution, representation of diagnoses, groups of patients.

Report
Epidemiology of cancer- trend
Epidemiology of cancer - selected diagnosis
Epidemiology of cancer - list of diagnoses

Report
Gender - trend
Gender - selected diagnosis
Gender - list of diagnoses
Age - trend
Age - selected diagnosis
Age - list of diagnoses
Clinical stage - trend
Clinical stage - selected diagnosis
Clinical stage - list of diagnoses

Report
Number of records - trend
Number of records - selected diagnosis
Number of records - list of diagnoses
Gender - trend
Gender - selected diagnosis
Gender - list of diagnoses
Age - trend
Age - selected diagnosis
Age - list of diagnoses
Clinical stage - trend
Clinical stage - selected diagnosis
Clinical stage - list of diagnoses

Report
Patients of Hxx by residence - trend
Patients of Hxx by residence - selected diagnosis
Patients of Hxx by residence - list of diagnoses
Oncology load of Hxx - trend
Oncology load of Hxx - list of diagnoses
Clinical stage - trend
Clinical stage - selected diagnosis
Clinical stage - list of diagnoses

Reporting will be designed as a set of dynamic reports

MU IBA

I-COP Cancer Care Information Center

Home | Reporting | Browser | Analyses | Help | Info | log in

Displayed units

Displayed value:
- incidence
- mortality
- inc. & mor.
- M/I index
- prevalence

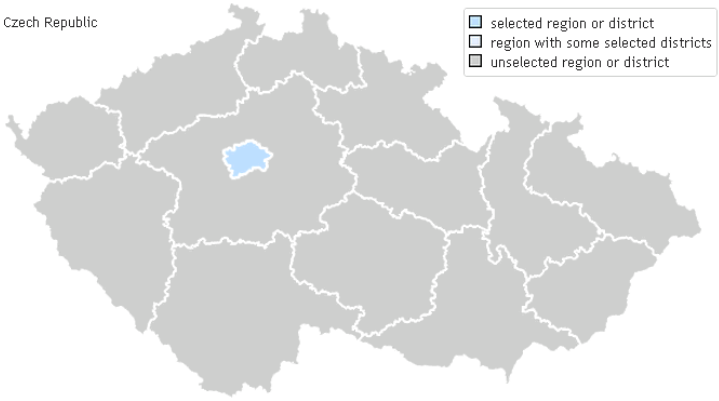
Displayed units:
- absolute numbers
- % share
- conversion to 100 000
- ASR (E)
- ASR (W)

Aggregate functions:
- sum
- average
- median

Study population

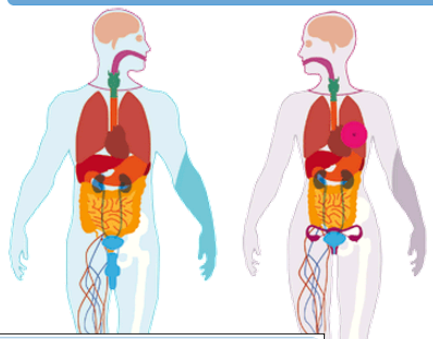
Whole country | Regions | Districts

- study population is from selected (blue) regions



Show

Diagnosis



Group of diagnoses: Breast - female (C50)

Selected diagnoses: C50 - female breast cancer

Show

Age

0 - 85+

Show

Stage of diagnosis according to NOR

Stage of tumor
- stage I
- stage II
- stage III
- stage IV
- N/A - objective reasons
- N/A - incomplete record

Show

Analysis selection:

Analysis setup:
- Study population
- Reference population
- Displayed units

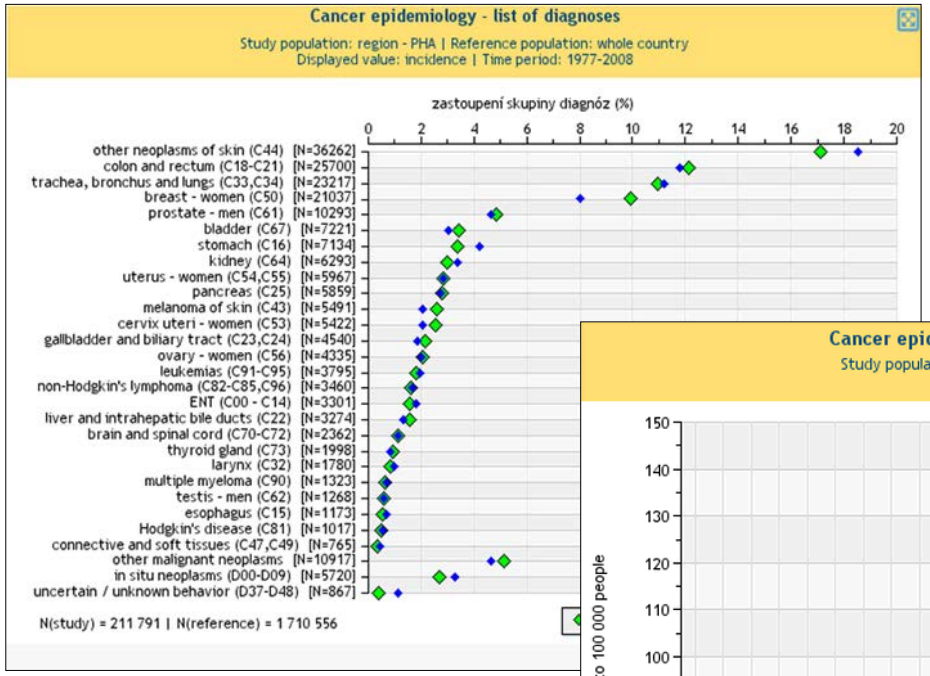
Analysis filter:
- Diagnosis
- Gender
- Age
- Period
- Stage
- T,N,M

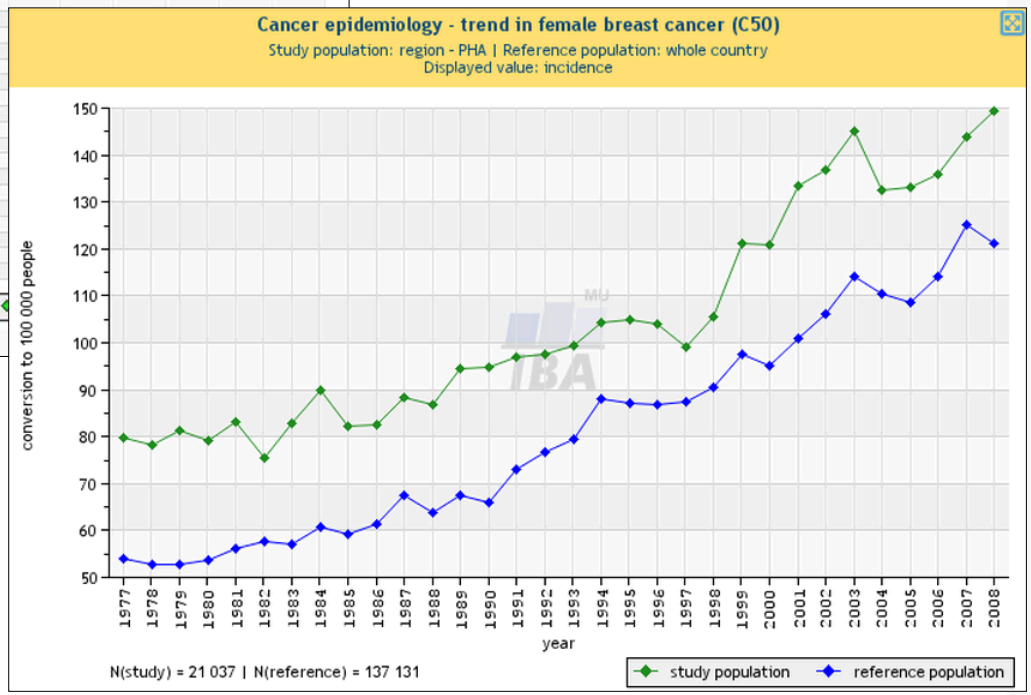
Other options:
- Order

I-COP Reporting will support options/filters to customize displayed reports

I-COP Cancer Care Information Center

Home | Reporting | Browser | Analyses | Help | Info | log in





Analysis selection:

Analysis setup:

- Study population
- Reference population
- Displayed units

Analysis filter:

- Diagnosis
- Gender
- Age
- Period
- Stage
- T,N,M

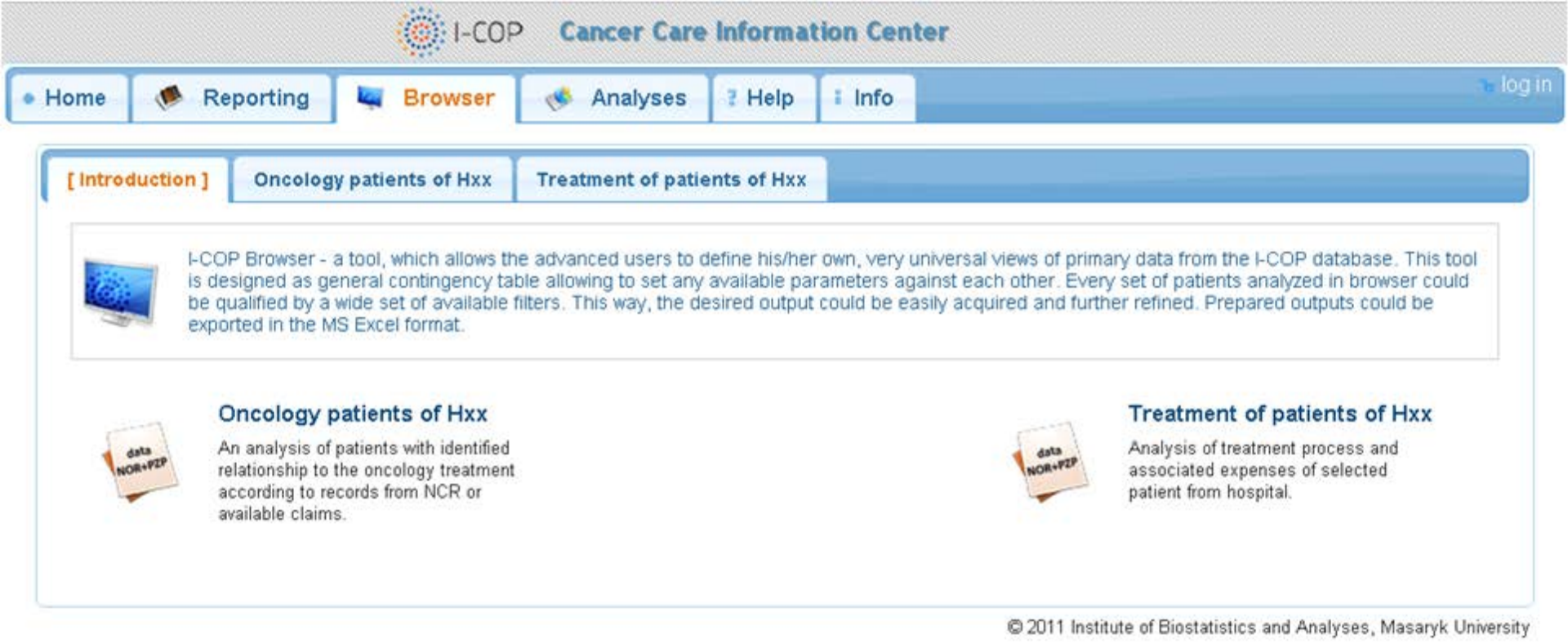
Other options:

- Order

Predefined customizable reports will offer the analyses of the processed hospital data. Comparing the specific hospital results with national reference values will be available to evaluate the hospital outcomes.

I-COP is designed as a tool for information support of physicians caring for cancer patients. Based on data fusion project of the National Cancer Registry and operational-administrative data of the payer health care. Its main ambition is to create an information base that will provide a combined output of these sources and define further analysis and their implementation.

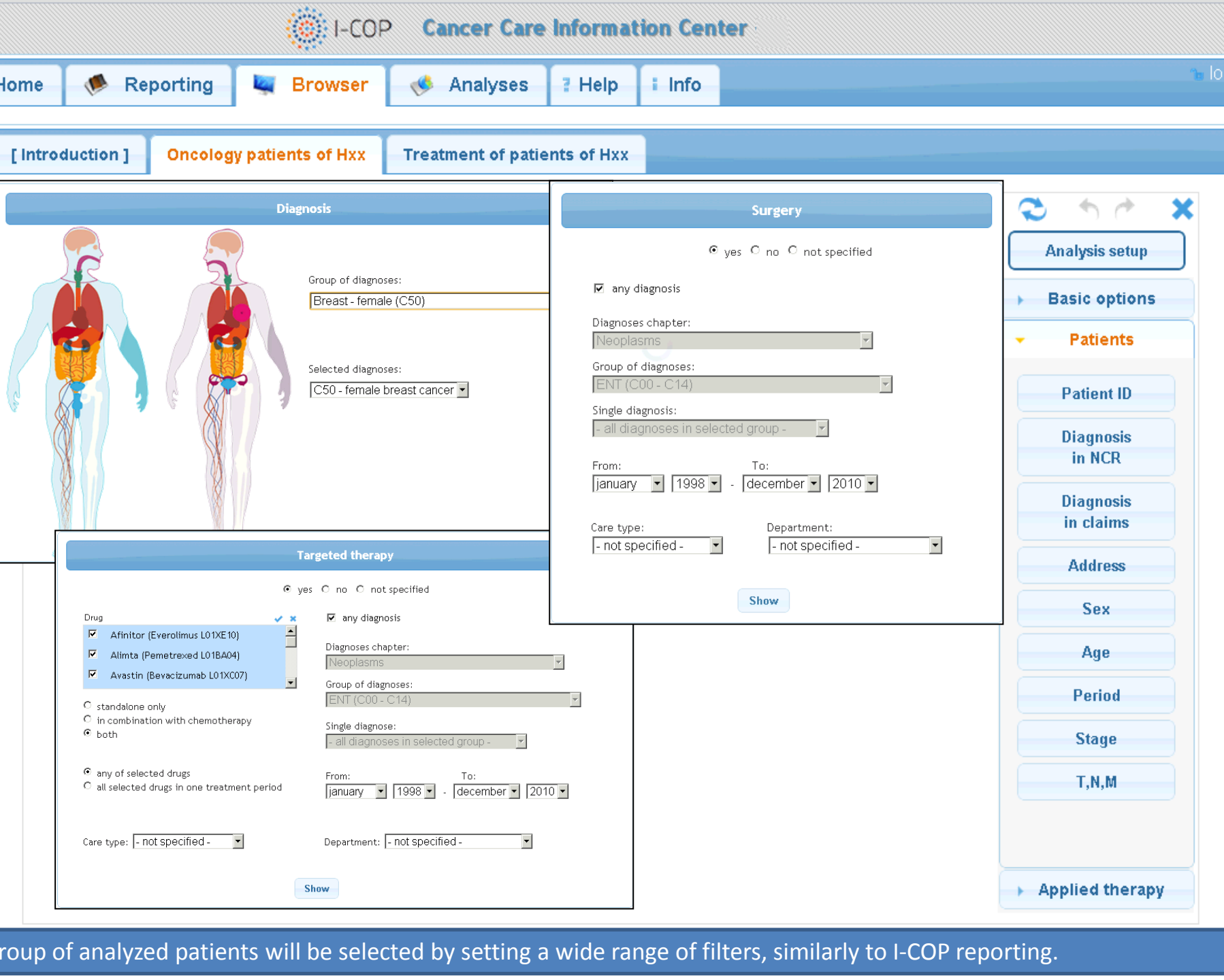
I-COP BROWSER



Report name	Description
A: Oncology patients treated in Hxx	Group of patients, for which the records of tumor diagnostics in NCR have been acquired.
B: Patients of Hxx without tumor diagnosis confirmation	Patients with possible oncology treatment in Hxx, but without any tumor record in NCR.
X: Oncology patients of Hxx without recorded treatment	Oncology patients with record assigned to Hxx without any claims in Hxx data.

Report name	Description
Treatment overview of selected patient	Overview of procedures and associated expenses during treatment of selected patient.

I-COP Reporting and I-COP Browser will be both integrated in the same environment of proposed web based application.



Group of analyzed patients will be selected by setting a wide range of filters, similarly to I-COP reporting.

I-COP Cancer Care Information Center

Home | Reporting | Browser | Analyses | Help | Info | log in

[Introduction] | Oncology patients of Hxx | Treatment of patients of Hxx

Oncology patients treated in Hxx (Patients: 361 of 134021)

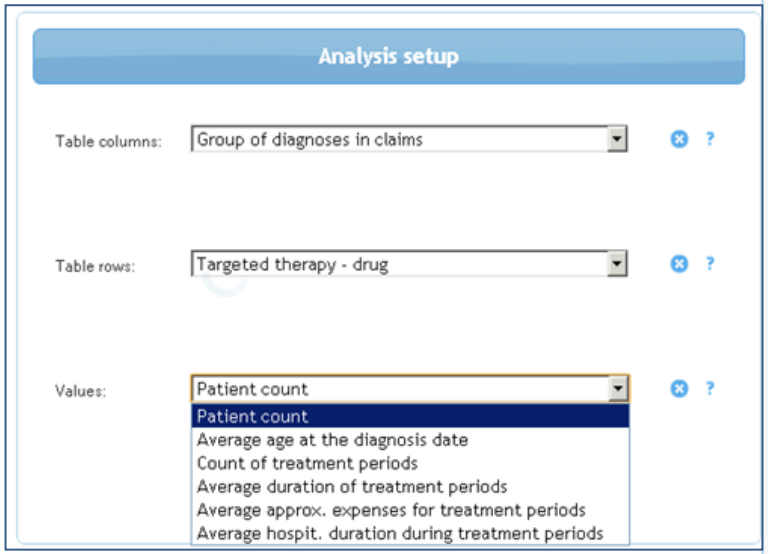
Basic options - without limitation
Patients - DG. in claims: female breast cancer (C50)
Applied therapy - Targeted: Erbitux,Herceptin,...

table rows: Targeted therapy - drug | table columns: Group of diagnoses in claims

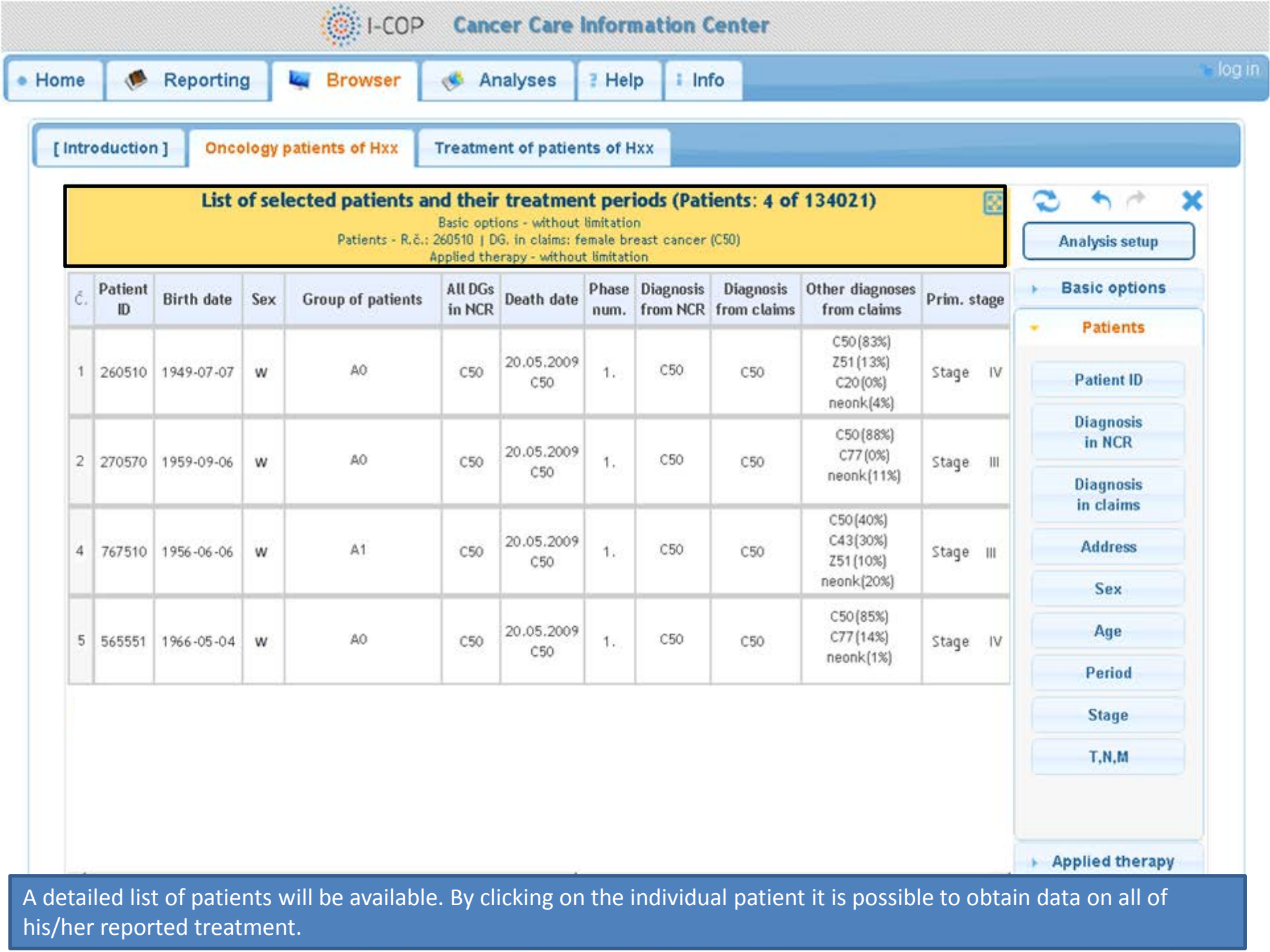
	Value	breast - women (C50)
Erbitux - Cetuximab	Patient count	3
	Proportion	0.8%
Herceptin - Trastuzumab	Patient count	358
	Proportion	99.2%
Tyverb - Lapatinib	Patient count	18
	Proportion	5%

Analysis setup

Basic options
Patients
Applied therapy
Care type
Surgery
Radiotherapy
Chemotherapy
Targeted therapy
Hormonal therapy
Imunotherapy
Bisfosfonates



The main output will be an interactive pivot table, where clicking on a single table cell leads to a list of patients fulfilling the inclusion criteria of this cell.



Č.	Patient ID	Birth date	Sex	Group of patients	All DGs in NCR	Death date	Phase num.	Diagnosis from NCR	Diagnosis from claims	Other diagnoses from claims	Prim. stage
1	260510	1949-07-07	W	A0	C50	20.05.2009 C50	1.	C50	C50	C50(83%) Z51(13%) C20(0%) neonk(4%)	Stage IV
2	270570	1959-09-06	W	A0	C50	20.05.2009 C50	1.	C50	C50	C50(88%) C77(0%) neonk(11%)	Stage III
4	767510	1956-06-06	W	A1	C50	20.05.2009 C50	1.	C50	C50	C50(40%) C43(30%) Z51(10%) neonk(20%)	Stage III
5	565551	1966-05-04	W	A0	C50	20.05.2009 C50	1.	C50	C50	C50(85%) C77(14%) neonk(1%)	Stage IV

A detailed list of patients will be available. By clicking on the individual patient it is possible to obtain data on all of his/her reported treatment.

Detail of patient with ID 260510

DG in NCR: C509 (st. IV, 2008) | Woman | Born 1949 | Death 2009 | Period of treat.: whole treatment | Pacienti: A0 - onco patients of Hxx

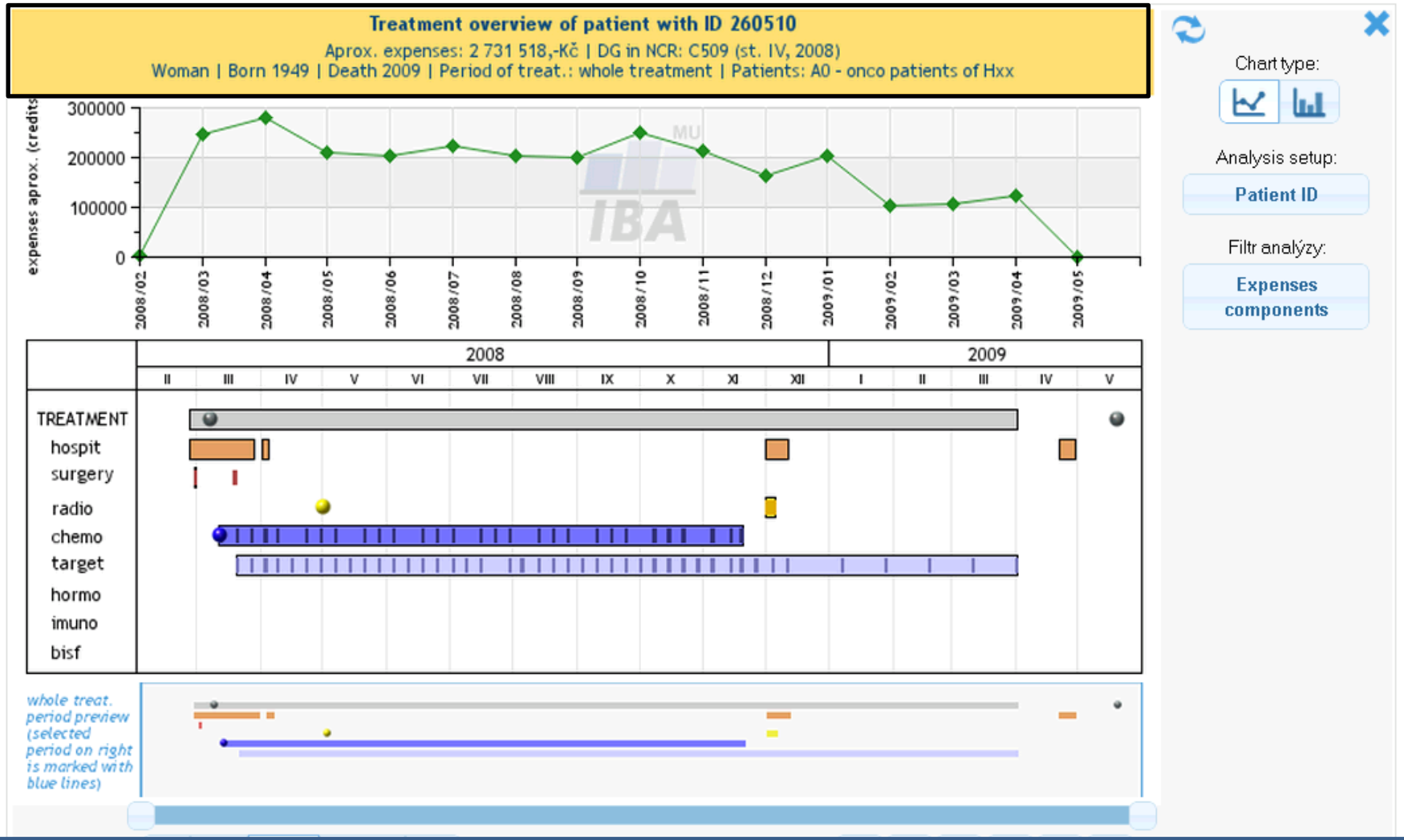
Č.	Treatment	DG of treat.	Record type	Date	Modality	Item code	Item name	Amount	Approx. costs
445	Primární léčba (27.02.2008-01.04.2009)	C50 ZN prsu	Ambulantní ošetření C50	2008-04-09	ostatní	0001203	Příplatek za ředění cytostatik v lékárně	1	256
446	Primární léčba (27.02.2008-01.04.2009)	C50 ZN prsu	Ambulantní ošetření C50	2008-04-09	ostatní	09220	KANYLACE PERIFERNÍ ŽÍLY VČETNĚ INFÚZE	1	149
447	Primární léčba (27.02.2008-01.04.2009)	C50 ZN prsu	Ambulantní ošetření C50	2008-04-09	ostatní	42510	NÁROČNÁ APLIKACE REŽIMŮ LÉČBY CYTOSTATIKY (1 DEN, NEZAHRNUJE PŘÍPRAVU LÉČIV)	1	451
448	Primární léčba (27.02.2008-01.04.2009)	C50 ZN prsu	Ambulantní ošetření C50	2008-04-09	ostatní	42520	APLIKACE PROTINÁDOROVÉ CHEMOTERAPIE	2	198
449	Primární léčba (27.02.2008-01.04.2009)	C50 ZN prsu	Ambulantní ošetření C50	2008-04-09	TARGET	0025555	HERCEPTIN 150 MG (Trastuzumab L01XC03)	2	37 803
450	Primární léčba (27.02.2008-01.04.2009)	C50 ZN prsu	Lab. vyšetření C509	2008-04-16	LAB		P_Kreatinin: 193.91 (mír.zvýš.)	0	0
451	Primární léčba (27.02.2008-01.04.2009)	C50 ZN prsu	Lab. vyšetření C509	2008-04-16	LAB		P_Natrium: 134.63 (v normě)	0	0
452	Primární léčba (27.02.2008-01.04.2009)	C50 ZN prsu	Lab. vyšetření C509	2008-04-16	LAB		P_Albumin: 30.62 (mír.sníž)	0	0
453	Primární léčba (27.02.2008-01.04.2009)	C50 ZN prsu	Lab. vyšetření C509	2008-04-16	LAB		P_Glukóza: 7.16 (mír.zvýš.)	0	0
454	Primární léčba (27.02.2008-01.04.2009)	C50 ZN prsu	Lab. vyšetření C509	2008-04-16	LAB		P_Urea: 13.69 (mír.zvýš.)	0	0
455	Primární léčba (27.02.2008-01.04.2009)	C50 ZN prsu	Ambulantní ošetření C50	2008-04-16	ostatní	09543	REGULAČNÍ POPLATEK ZA NÁVŠTĚVU -- POPLATEK UHRAZEN	1	0
456	Primární léčba (27.02.2008-01.04.2009)	C50 ZN prsu	Ambulantní ošetření C50	2008-04-16	ostatní	09119	ODBĚR KRVE ZE ŽÍLY U DOSPĚLÉHO NEBO DÍTĚTE NAD 10 LET	1	24
457	Primární léčba (27.02.2008-01.04.2009)	C50 ZN prsu	Ambulantní ošetření C50	2008-04-16	ostatní	42023	KONTROLNÍ VYŠETŘENÍ ONKOLOGEM	1	159
458	Primární léčba (27.02.2008-01.04.2009)	C50 ZN prsu	Ambulantní ošetření C50	2008-04-16	ostatní	81329	ALBUMIN (SÉRUM)	1	15

A detailed list of all the reported data sorted by procedure date and the type of treatment will be available. This output will display the primary raw data.

I-COP Cancer Care Information Center

Home | Reporting | Browser | Analyses | Help | Info | log in

[Introduction] | Oncology patients of Hxx | Treatment of patients of Hxx



Last proposed outcome is a graph providing an overview of the applied cancer treatment procedures. The treatment scheme of individual patient could be easily evaluated by oncologists.

Part III

Using I-COP application as an information support for the screening programs in hospitals (colorectal carcinoma screening, short example)

Intra-hospital example

- We want to know:
  - How many screening colonoscopies are performed in our hospital?
  - How many patients with colonoscopy were positive-diagnosed in our hospital?
  - Can we learn more about the positive-diagnosed patients?

I-COP Cancer Care Information Center

Home | Reporting | Browser | Analyses | Help | Info | log in

[Introduction] | Oncology patients of Hxx | Treatment of patients of Hxx

Oncology patients treated in Hxx (Patients 4019 of 134021)

Basic options - without limitation
Patients - DG. in claims: without limitation
Applied therapy - 15404,15105,...

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Patient count	725	803	876	974	1122	1250	1473	1616	1767	1984	2006	2038	1979
Proportion	18%	20%	21.8%	24.2%	27.9%	31.1%	36.7%	40.2%	44%	49.4%	49.9%	50.7%	49.2%

Analysis setup

- Basic options
- Patients
- Applied therapy
  - Care type
  - Surgery
  - Radiotherapy
  - Chemotherapy
  - Targeted therapy
  - Hormonal therapy
  - Imunotherapy
  - Bisfosfonates

Applied medicaments | Medical procedures

yes / no

Search medical procedures: 15101

Search and sort by: name / code

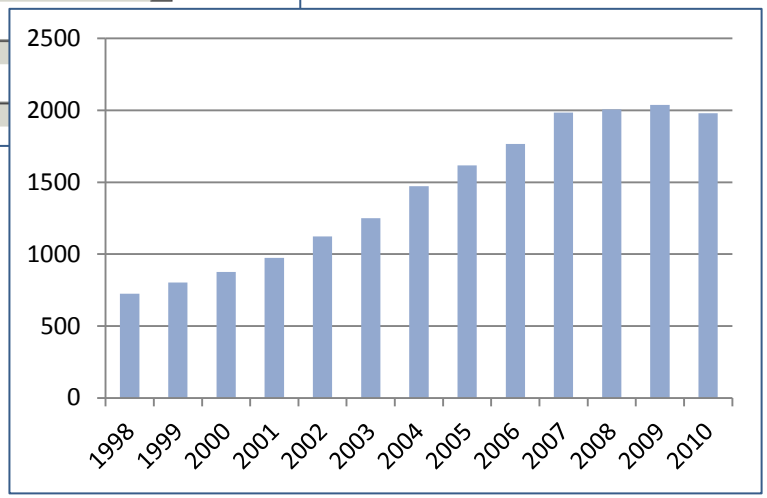
Selected medical procedures: SCREENING COLONOSCOPY

any diagnosis

Diagnosis chapter: Neoplasms

Group of diagnosis: ENT (C00 - C14)

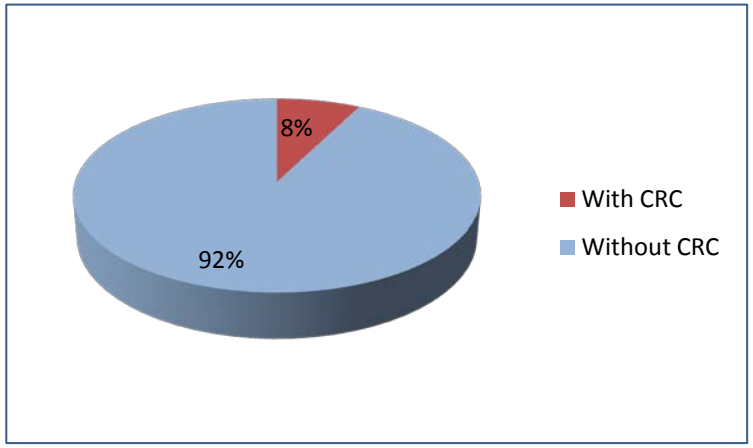
Single diagnosis: - all diagnoses in group -

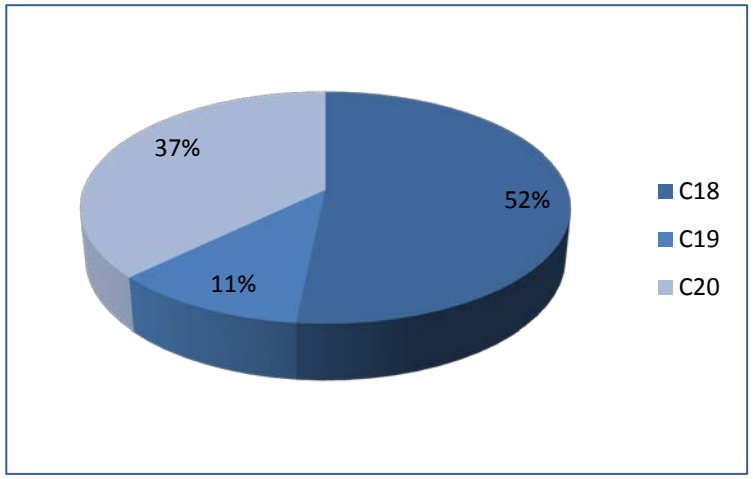


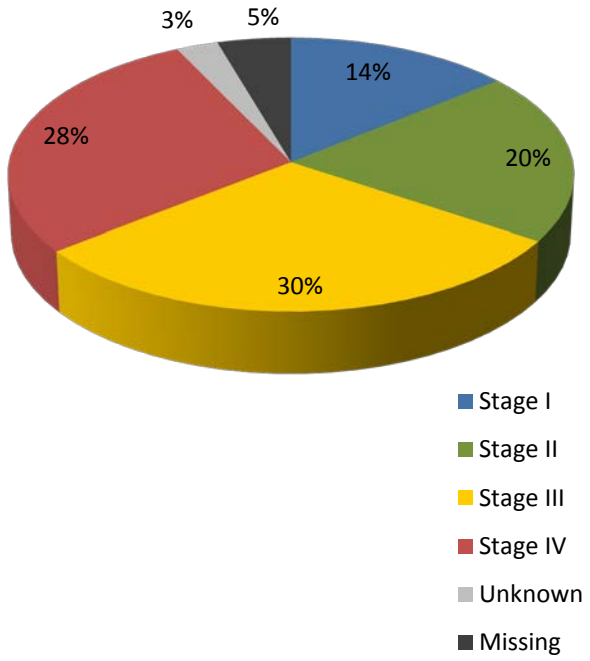
We can select all the patients with screening colonoscopy performed and divide this group by the year of the procedure, as we can see in the pivot table.

Oncology patients treated in Hxx (Patients 4019 of 134021)

Basic options - without limitation
Patients - DG. in claims: without limitation
Applied therapy - 15404,15105,...







Analysis setup

- Basic options
- Patients
- Applied therapy
  - Care type
  - Surgery
  - Radiotherapy
  - Chemotherapy
  - Targeted therapy
  - Hormonal therapy
  - Imunotherapy
  - Bisfosfonates

Now we are able to analyze the selected group of patients.

Conclusion

- We want to design and develop a flexible educational application, which will be able to process, visualize and analyze the real life hospital data.
- Because most of the CCCs centers participate on this project, we will be able to build up a really robust and representative data source for information support of the oncologists.
- This proposed application will have a potential to be a part of the screening programs as a intra-hospital screening monitor and a data source for the validating data in the screening registries.

KONEC