



Institute of Biostatistics and analyses: history and activities



Jiří Jarkovský, Adam Svobodník



Institute of Biostatistics and Analyses

History
Structure
Collaboration

Institute of Biostatistic and Analyses: History

☑ Since 2001

- group of scientists interested in analysis of biological and clinical data
- Centre of Biostatistics and analyses

☑ 2006

- Institute of Biostatistics and Analyses
- 45 employees and PhD students
- 40 students in pregraduate study of Computational Biology

☑ 2007

- Integration into Research Centre for Environmental Chemistry and Ecotoxicology



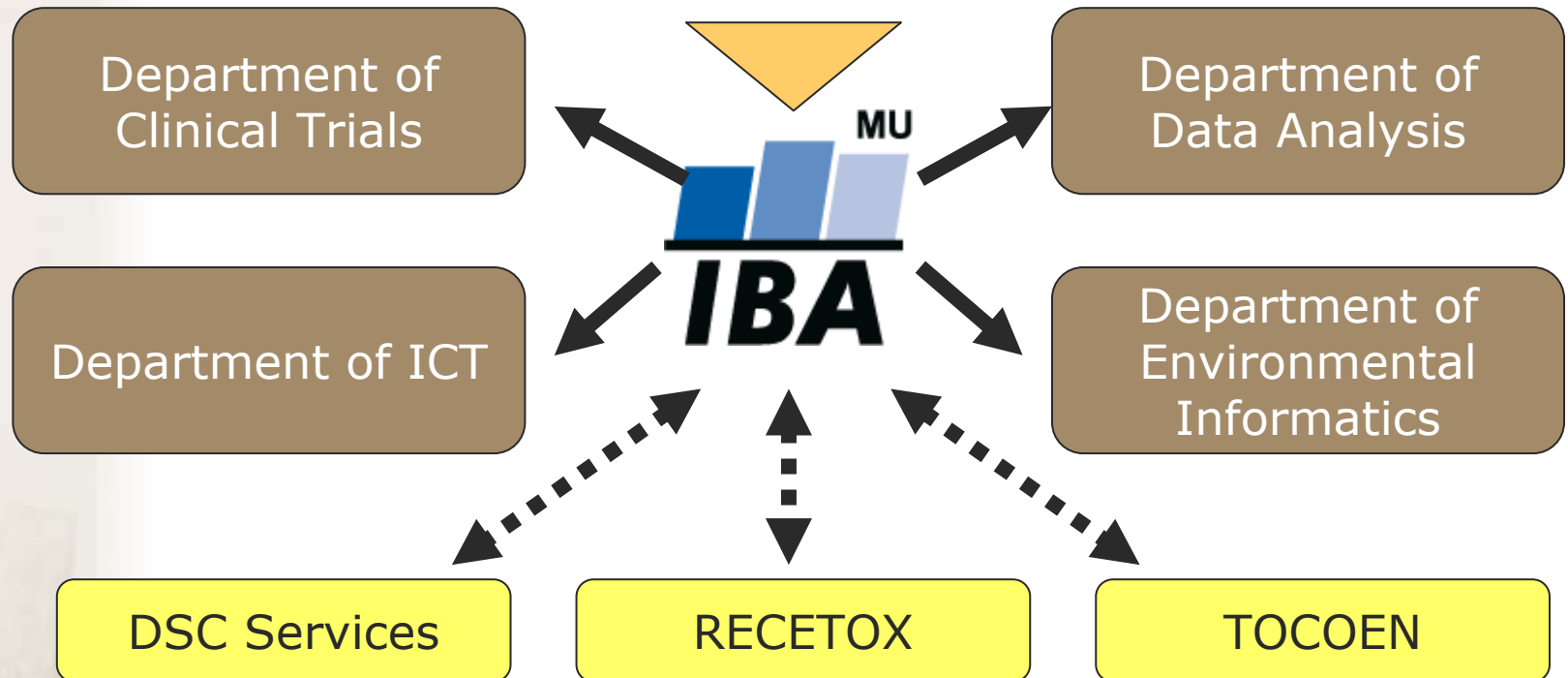
Institute of Biostatistic and Analyses: Internal Structure

✓ Nowadays IBA has the following structure:



Masaryk university

- Faculty of Medicine
- Faculty of Science



Institute of Biostatistic and Analyses: Collaboration

**Close mutual
collaboration**

- Faculty of Medicine (MU in Brno)
- Faculty of Informatics (MU in Brno)
- Faculty of Science (MU in Brno)
- Research Centre for Environmental Chemistry and Ecotoxicology (MU in Brno)
- Department of Mathematics (MU in Brno)

- Czech Environmental Information Agency
- National and Regional Centre for Persistent Organic Pollutants in Middle and East Europe
- Masaryk Memorial Hospital
- Czech Statistical Office
- Czech National Cancer Registry
- Committee for Breast Cancer Screening (Ministry of Health)
- Czech Oncology Society

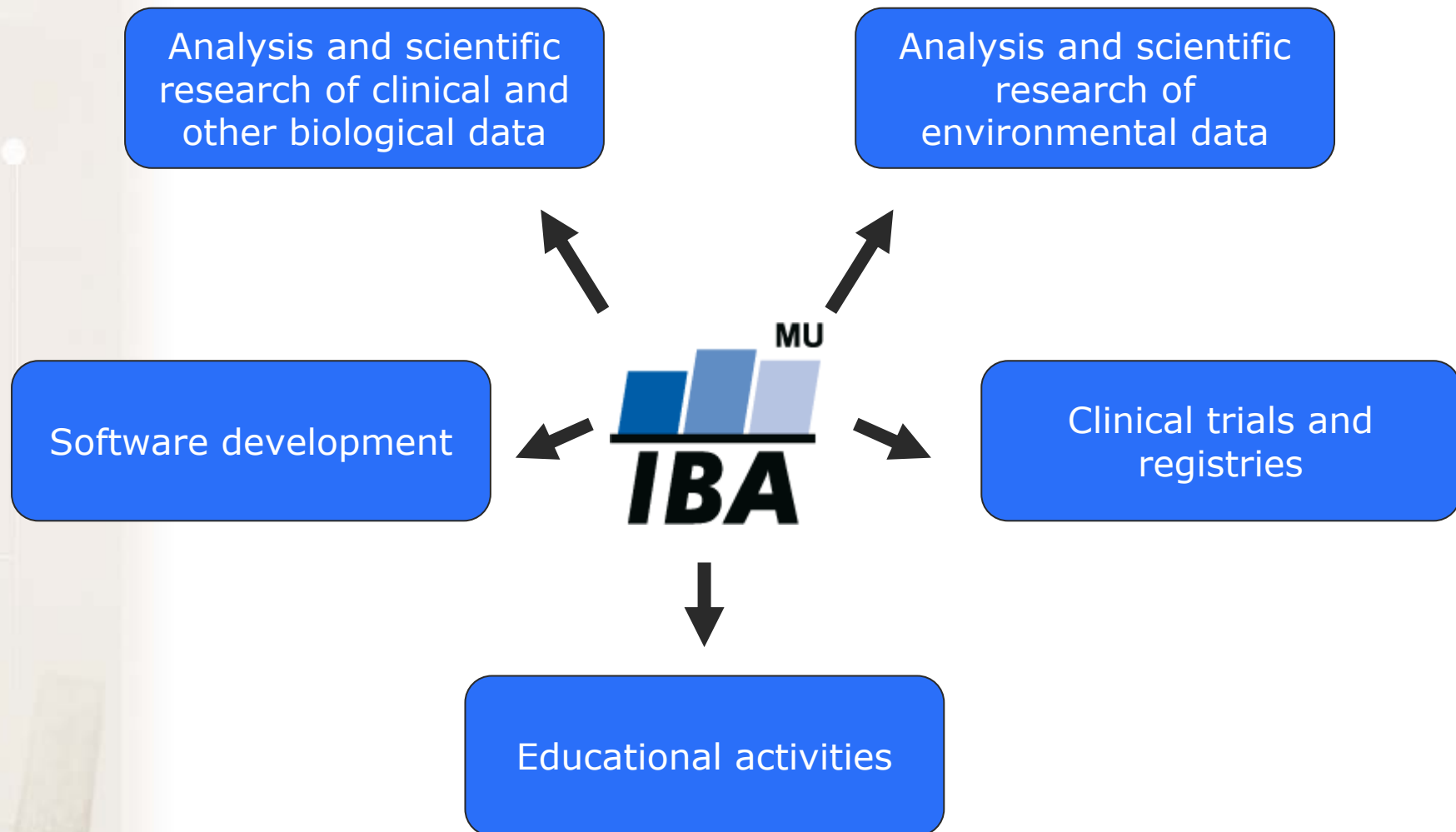


**Standard professional
collaboration**

More than 100 independent health care providers and hospitals, research institutions, environmental agencies, ministries etc.

Institute of Biostatistic and Analyses: Areas of interests

- ☑ There are five main areas of interest



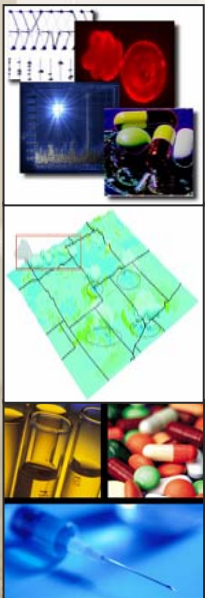


Educational activities

Basic and advanced courses of data analysis
Computational Biology
Conferences
Background of educational and scientific activities

Computational biology

- ✓ **Computational biology is newly (since 2001) established study programme at Masaryk University with high potential in almost all biological and clinical disciplines**
- ✓ **The concept of the computational biology is:**
 - intrinsically inter-disciplinary
 - closely related to standard biological and clinical disciplines and their modern trends (bioinformatics, processing of large scale data from genomics to ecological biomonitoring, risk assessment, predictive modeling...).
- ✓ **We define computational biology in a broad sense as umbrella branch for all standard biological and clinical disciplines with three main branches:**



□ Analysis of biological and clinical data

- Applied analysis of data or development of new methods

□ Environmental informatics

- Environmental information systems, biomonitoring etc.

□ Informatics in medical sciences

- Informatics in clinical trials, information systems for clinical data, clinical registry

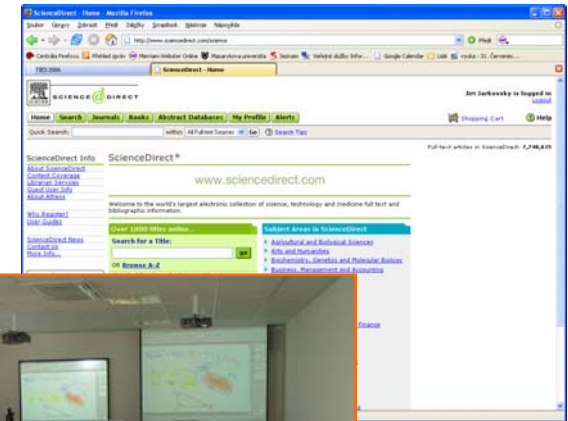
Conferences

- ☑ **IBA organizes several international scientific meeting and conferences**
 - **2005**
 - ☐ **Summer School of Environmental Chemistry and Ecotoxicology 2005: tutorial of environmental risk assessment**
 - ☐ **Enviroinfo 2005 – Informatics for environmental protection**
 - ☐ **Summer School on Computational Biology**
 - **2006**
 - ☐ **Summer School of Environmental Chemistry and Ecotoxicology 2006: tutorial of environmental risk assessment**
 - ☐ **Summer School on Computational Biology: Predictive modelling and ICT in Environmental Research – September 11-13**
 - **2007**
 - ☐ **TIES: Conference of International Environmetrics Society**
 - ☐ **Summer School on Computational Biology: Processing and Analysis of Biodiversity Data: from Genomic Diversity to Ecosystem Structure – August 13-15**

Background of educational and scientific activities

✓ Library

- Books and journals covering almost all areas of data analysis in biology and medicine
- Free access to many databases



✓ Software equipment

- Statistica
- SPSS + Clementine
- SAS
- R
- Matlab
- Maple
- ArcGIS



✓ Modern classroom

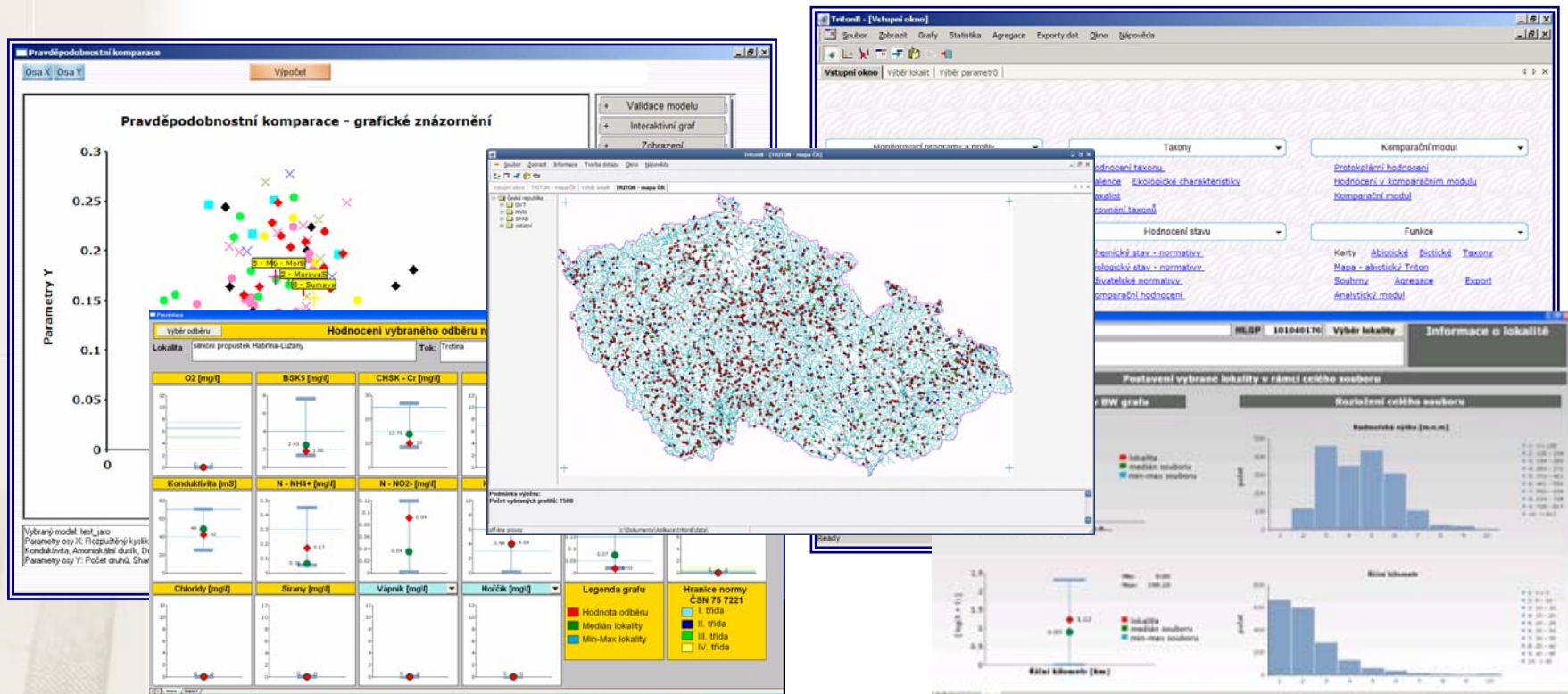


Important projects in environmental, biological and epidemiology research

Environmental research
Epidemiology
Various biological research

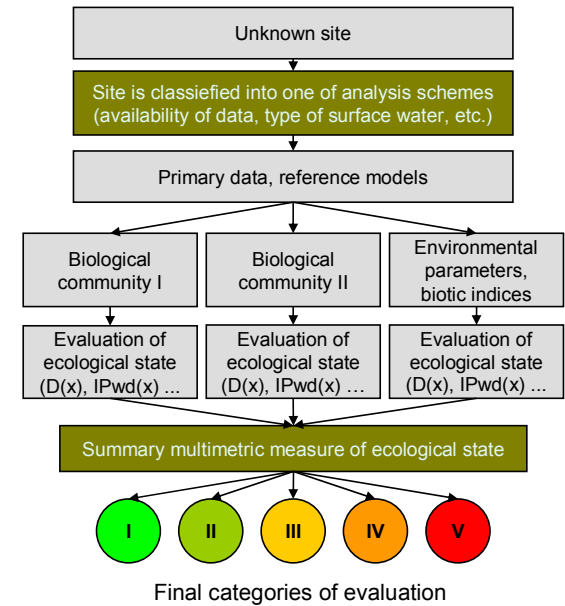
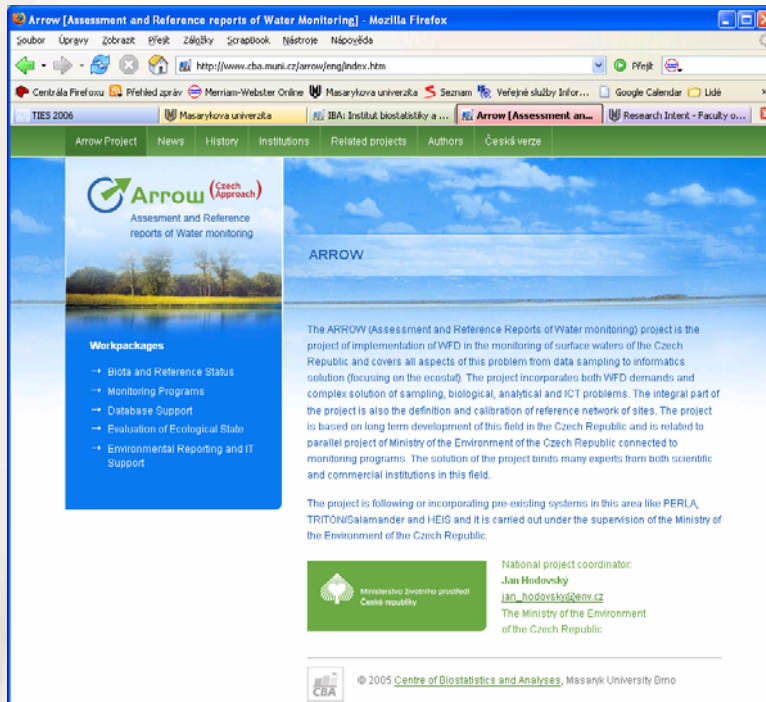
TRITON project

- ✓ Project with Agricultural Water Management Authority
- ✓ Software and statistical methodology of data analysis in biomonitoring networks
 - ➔ Visualisation of data
 - ➔ Statistical analysis of time trends, biodiversity etc.



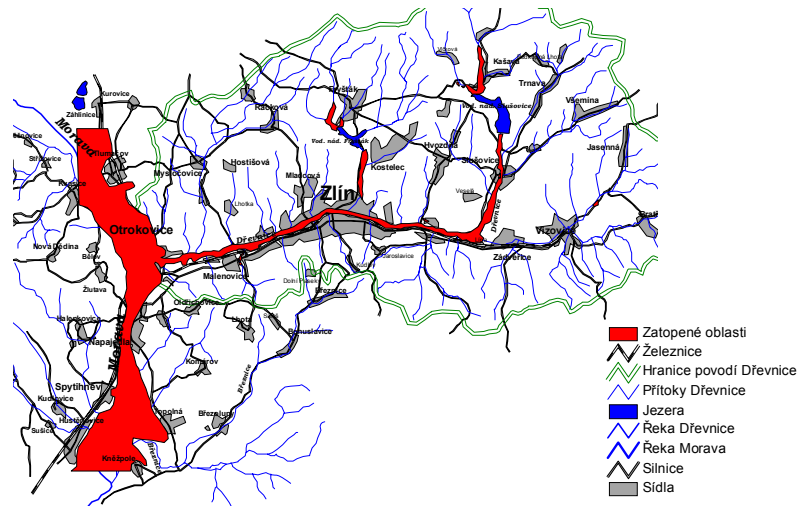
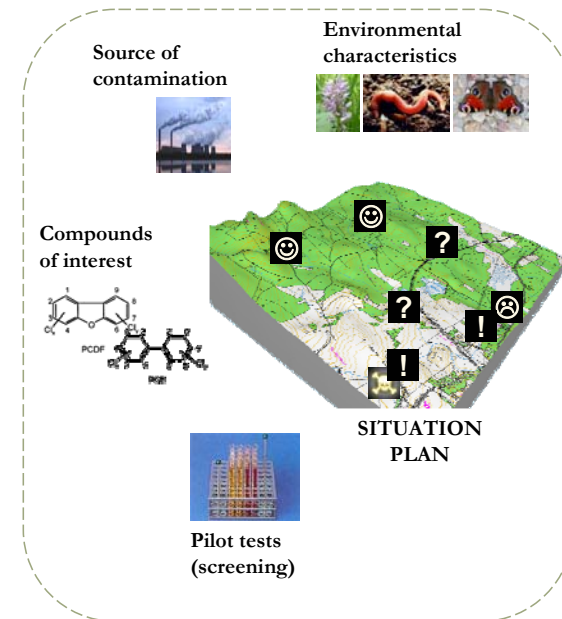
ARROW project

- ☑ Part of implementation of Water Framework Directive EU in the Czech Republic
- ☑ IBA tasks
 - ➔ Development of methodology and scientific research in analysis, typology and prediction of ecological state of localities
 - ➔ Software implementation of analytical methodology



INCHEMBIOL project

- ☑ INTERACTIONS AMONG THE CHEMICALS, ENVIRONMENT AND BIOLOGICAL SYSTEMS AND THEIR CONSEQUENCES ON THE GLOBAL, REGIONAL AND LOCAL SCALES
- ☑ Long term projects (7 years) including
 - ➔ Experimental work
 - ➔ Data analysis
 - ☐ Modelling of POPs distribution and transport
 - ☐ Dose-response analysis
 - ☐ Biodiversity analysis
 - ☐ Analysis of data of biomonitoring networks
 - ☐ Environmental risk assessment
 - ☐ GIS



INCHEMBIOL: Environmental risk assessment



Uncertainty analysis



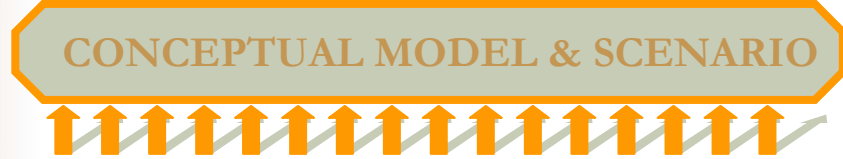
Benchmarking
Probability estimation



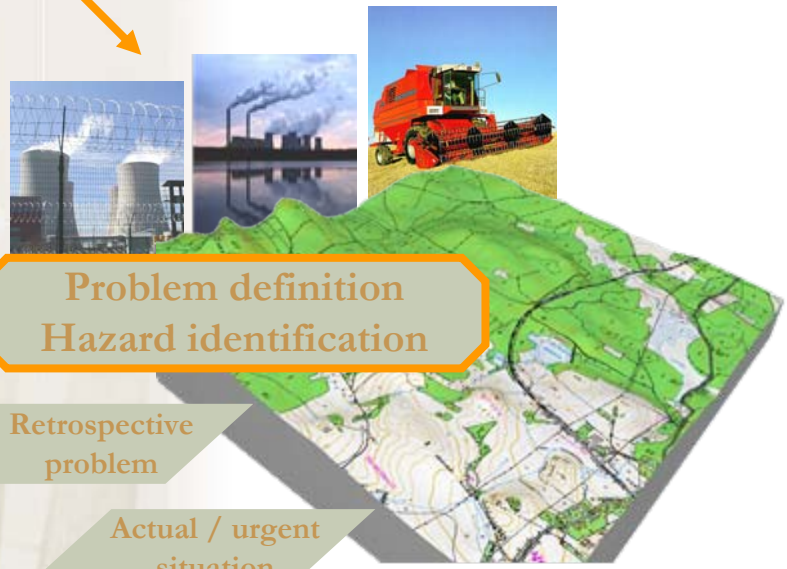
Bio-tests Bioindicators Biomonitoring



Data processing Modelling
Multivariate analyses



Model of area
of interest
Experimental design



**Problem definition
Hazard identification**

Retrospective
problem

Actual / urgent
situation

Prospective
evaluation

Data gathering
Data aggregation
Information services

Optimization
&
Processing
in information
systems
&
Communication

Jarkovský, Sv

SVOD project: description

- ✓ **SVOD = EXPERT SYSTEM FOR ANALYSIS AND PRESENTATION OF EPIDEMIOLOGICAL DATA IN ONCOLOGY**
- ✓ **Data of National Cancer Institute: population database of cancer epidemiology since 1970**
- ✓ **Data includes incidence and mortality rates, time profiles, distribution of clinical stages, regional heterogeneity and more...**

http://www.svod.cz - SVOD Analyse - Mozilla Firefox

REGIONAL OVERVIEWS

Select a diagnose

II. TUMOURS OF DIGESTIVE ORGANS

- C09,C10,C12-C14 - Malignant neoplasm of other sites in the pharynx
- C15 - Malignant neoplasm of esophagus
- C16 - Malignant neoplasm of stomach
- C17 - Malignant neoplasm of small intestine
- C18 - Malignant neoplasm of colon
- C19 - Malignant neoplasm of rectosigmoid junction
- C20 - Malignant neoplasm of rectum
- C21 - Malignant neoplasm of anus and anal canal
- C18-C21 - Malignant neoplasm of colon and rectum
- C22 - Malignant neoplasm of liver and intrahepatic bile ducts
- C23 - Malignant neoplasm of gallbladder
- C24 - Malignant neoplasm of other and unspecified parts of biliary tract
- C23,C24 - Malignant neoplasm of gallbladder and biliary tract
- C25 - Malignant neoplasm of pancreas
- C26 - Malignant neoplasm of other and ill-defined digestive organs

Display all diagnostic groups

http://www.svod.cz/analyse.php?modul=regionprehled#



http://www.svod.cz - SVOD Analyse - Mozilla Firefox

TIME TRENDS

Patient group selection

Sex, age Area Period

Stage TNM classification Other parameters

Other analyses for selected diagnose and group of patients

C50,D05 - Breast tumours, women
growth index of incidence to year 1977

Number of cases

3500
3000
2500
2000
1500
1000
500
0
-500

1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003

Analysed data: N=107294 http://www.svod.cz Source of data: ÚZIS

Other outputs

Change diagnose Analysis settings Data table Display report

Hokovo

SVOD project: aims

✓ Project aims

➔ Scientific research of epidemiology data

- ❑ General epidemiological and socio-economical analyses.
- ❑ Load of population by malignant tumours, population risks.
- ❑ Analysis of health care heterogeneity
- ❑ Approximate evaluation of the health care effectiveness according to therapeutic results.

➔ Software development

- ❑ Desktop software
- ❑ Web visualization of epidemiology data – public access

<http://www.svod.cz/>

SVOD - Mozilla Firefox

Soubor Úpravy Zobrazení Efekty Záložky Scrapbook Nástroje Nápověda

http://www.svod.cz/sec=@ualty@lang=en

Centrála Firefoxu Přehled zpráv Meriam-Webster Online Masarykova univerzita Seznam Všechny služby Infor... Google Calendar

TIRES 2006 Masarykova univerzita IBA: Institut biostatistiky... CENA SVOD

EPIDEMIOLOGY OF MALIGNANT TUMORS IN THE CZECH REPUBLIC

ABOUT PROJECT
NEWS
SOFTWARE SVOD
EPIDEMIOLOGICAL ANALYSES
ANALYSES WIZARD

Publications
Scientific events
WWW links

Other projects

Národní program mamografického screeningu v ČR
www.npmo.cz

Databáze pro management dat mamografického screeningu
www.cba.muni.cz/serodivness

LOAD USER

NEWS

Epidemiological analyses

INCIDENCE AND MORTALITY REGIONAL OVERVIEWS AGE OF PATIENTS

COMPARISON WITH FOREIGN COUNTRIES SUMMARY PRESENTATION Other analyses >>>

News

28.06.2005 Aktualizace dat MOR za rok 2003
V epidemiologických analýzách jsou nyní k dispozici data MOR za období 1977-2003.

10.10.2005 Výbor ČOS schválil odbornou radu projektu SVOD
V rámci odborních garancí projektu jsou zastoupeni představitelé všech významných směrů současné onkologie a odborní zástupci spolupracujících institucí.

04.10.2005 Probíhající úpravy portálu
Z důvodu validace dat jsou v analýzách dočasně nepřístupné diagnózy nádorů krevního a imunitního systému. Děkujeme za pochopení.

20.09.2005 Zprovoznění serveru
Dne 20.září 2005 byl tento webový portál zpřístupněn široké veřejnosti.

Press News from portal

Holovo

INFORMATION AND ANALYTIC SERVICES OF THE SYSTEM SVOD

SVOD kompletní verze - C64 Zhubný novotvar ledviny mimo pánevku

Soubor Nastavení Zobrazit Expertní služby nápověda

SVOD – System for Viewing of Oncological Data

CSO Malignant neoplasm of breast

Analytic tools – expert services

SVOD – Main menu

- Diagnose selection
- COBRA - Expert services
- User preferred analyses
- Web portal SVOD
- Discussion club SVOD
- SVOD manual
- SVOD help
- Exit

COBRA – Expert services

Presentations Data browser Expert services

User preferred analyses

- Epidemiology: incidence and mortality
- Comparative analysis: epidemiology
- Age structure of patient population
- Survival analysis
- Clinical stages
- Demographic data

Communication and information tools

Predefined user controlled analytic tools

Expert services

Expert services software SVOD

Epidemiology

Comparative analysis

Health care management

Základní epidemiologie

Základní epidemiologie
Diagnóza COO

Počet případů

Incidence, (n=4061)

Mortality, (n=717)

1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002

Equity Tabulka hodnot Zapiš

Interakční analýza pčezli

Analýza pčezli
Diagnóza CSO

Kumulativní podíl pčezlých

St. I, (n=1848)

St. II, (n=222)

Podíl případů s daným stádiem (%)

St. I, (n=1848)

St. II, (n=222)

St. III, (n=208)

St. IV, (n=11)

St. V, (n=1272)

1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002

Equity Tabulka hodnot Zapiš

Výběr specifické incidence

Výběr specifické incidence

Shmárovatý kraj Referenční hodnoty: maximum median minimum

Výběr region: Shmárovatý kraj
Výběr standard: všechny regiony v ČR Referenční hodnoty: kraj, median, minimum

Equity Tabulka hodnot Zapiš

Comprehensive presentations with comments and access to appropriate analytic tools

Vývoj věkové struktury v čase - všichni pacienti

Věk do 45 let

Věk 45 - 59 let

Věk 60 a více let

Dišní údaje o věkové struktuře pacientů s danou diagnózou a údaje o věkové struktuře populace můžete zhlédnout v podobě grafických a tabulkových výstupů pomocí těchto uživatelských nástrojů analýzy:

Analýza věkového složení Výběrové specifické analýzy v čase Demografické údaje

Data browser - analysis of individual parameters

Prohlížeč dat / tvorba výběru

Negativně analyzovaný parametr ze seznamu. Po výběru analyzovaného parametru můžete v ostatních záložkách běžně učit skutečné parametry ze seznamu chcete-li dále pracovat.

Analyzovaný parametr	Region	Období	Věková skupina	Pohlaví	Diagnóza	Klinická stádia	Další výběr
1 pohlaví	kategorie	2470					
2 rok narození	datum	2470					
3 věk při stanovení dg.	spojitý	2470					
4 věk při dg. - 5letá kateg.	kategorie	2470					
5 kraj trvalého bydliště	kategorie	2470					
6 sociální postavení	kategorie	374					
7 hlavní životní zaměstnání	kategorie	374					
8 ZN v rodinné anamnéze	kategorie	374					
9 kouření	kategorie	374					
10 diagnóza onemocnění	kategorie	2470					
11 rok stanovení diagnózy	datum	2470					
12 laterality	kategorie	2470					

základní parametry diagnostické parametry klinické parametry Účet

Výsledný graf - kategorie

Analýzovaný parametr: Klinické stádium

Stadium I 30.0%

Stadium II 30.0%

Stadium III 17.0%

Stadium IV 18.7%

Stadium V 4.3%

Statisitické údaje - tabulka hodnot

Stadium	Stadium	Průměr
1	Stadium I	30761
2	Stadium II	20628
3	Stadium III	17947
4	Stadium IV	12768
5	Stadium V	8871
6	Tabulka	300284

Klinické stádium

1 2 3 4 5

IBA RECO U

Microarrays and analysis of their data

- ✓ Promising technique of molecular biology, nevertheless with lots of problems on the field of data analysis
- ✓ Analysis of expression and CGH microarrays data
- ✓ Project in cooperation with Masaryk Memorial Hospital
 - ➔ Genomic profiling in prediction of response of radiochemotherapeutic treatment in colorectal carcinoma

