

Open source tools for clinical research in a diabetes clinic

Iztok Štötl

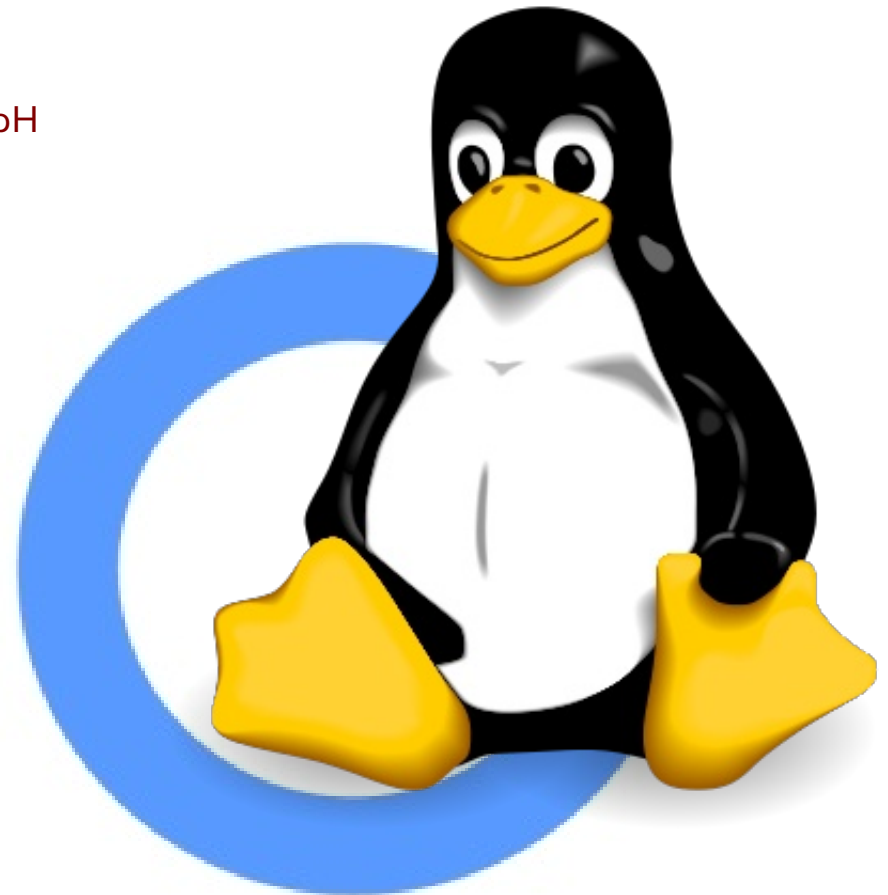
University Medical Centre Ljubljana

Peter Beck

JOANNEUM RESEARCH Forschungsgesellschaft mbH

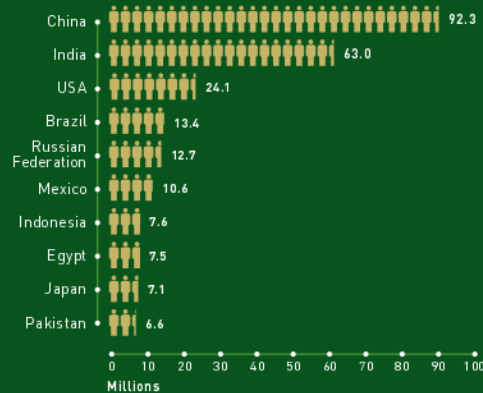


Berlin 2013



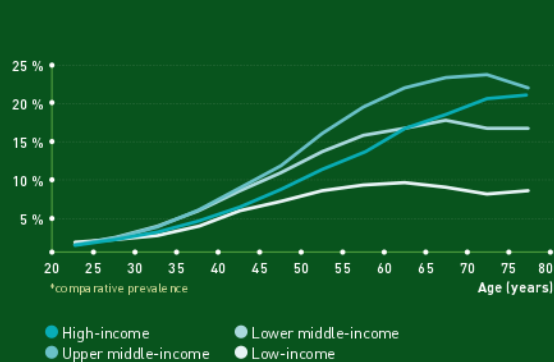
More than 371 million people have diabetes.

TOP 10 COUNTRIES/TERRITORIES FOR PEOPLE WITH DIABETES (20-79 YEARS)



4 out of 5 people with diabetes live in low- and middle-income countries.

PREVALENCE* (%) ESTIMATES OF DIABETES (20-79 YEARS) BY INCOME GROUP AND AGE



The number of people with diabetes is increasing in every country.

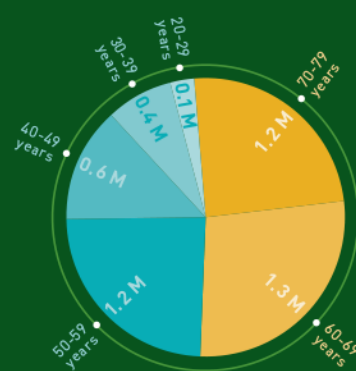
TOP 10 COUNTRIES/TERRITORIES FOR PREVALENCE* (%) OF DIABETES (20-79 YEARS)

COUNTRY /TERRITORY	PREVALENCE (%)
1 Federated States of Micronesia	37.2
2 Nauru	30.1
3 Marshall Islands	27.1
4 Kiribati	25.5
5 Tuvalu	24.8
6 Kuwait	23.9
7 Saudi Arabia	23.4
8 Qatar	23.3
9 Bahrain	22.4
10 Vanuatu	22.0

*comparative prevalence

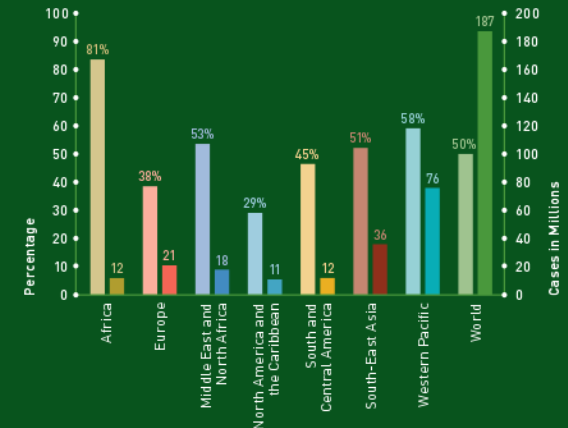
Half of people who die from diabetes are under the age of 60.

DEATHS ATTRIBUTABLE TO DIABETES BY AGE (20-79 YEARS)



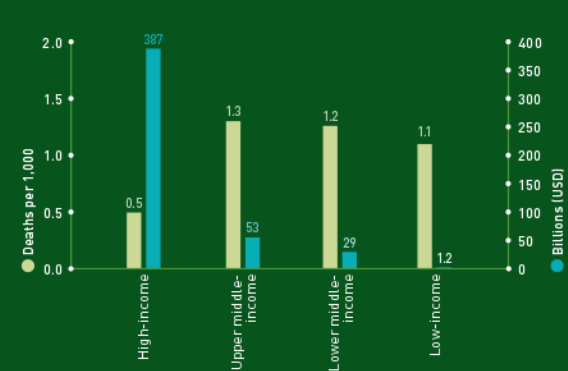
Half of people with diabetes don't know they have it.

UNDIAGNOSED PERCENTAGE AND UNDIAGNOSED CASES OF DIABETES (20-79 YEARS) BY REGION



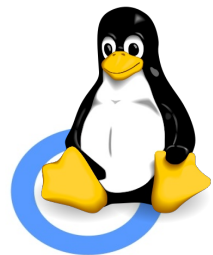
4.8 million people died and 471 billion USD were spent due to diabetes in 2012.

HEALTHCARE EXPENDITURES AND DEATHS PER 1,000 DUE TO DIABETES BY INCOME GROUP



Different IT problems in clinical diabetology

- Standardisation of Electronic Medical Record (still a lot of work to do)
- Hospital and ambulatory information systems (mostly proprietary and closed source)
- Telemedicine, PHR and different medical gadgets (insulin pumps, sensors ...)
- **Data management in Clinical research**
- **Computerised diabetes registries**



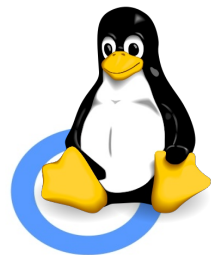


University medical centre Ljubljana, Department of Endocrinology, Diabetes and Metabolic Diseases

Small department with inpatient and outpatient
diabetes clinic

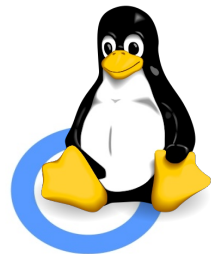
Own servers:

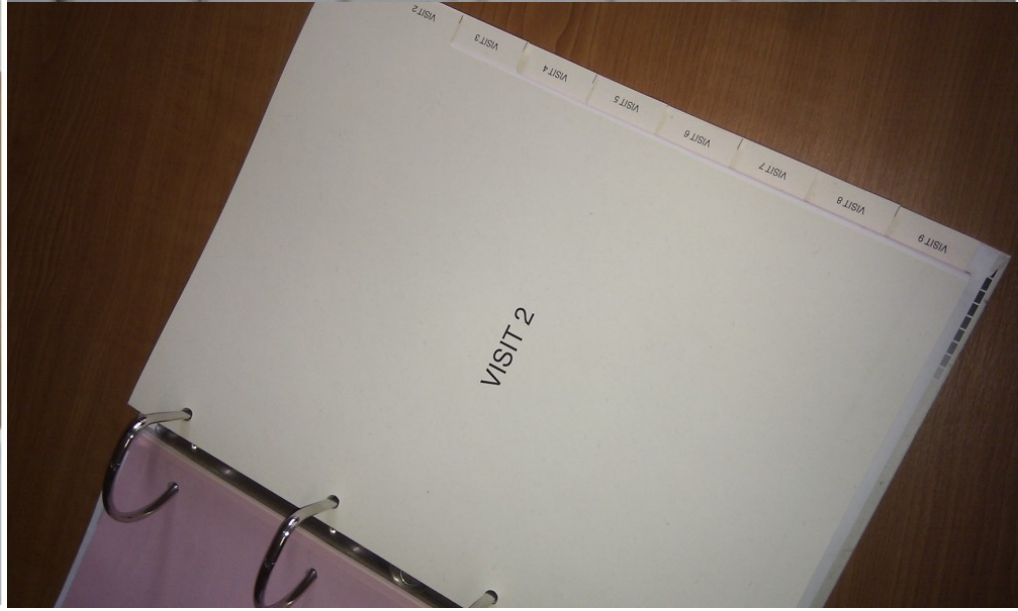
- Centos
- KVM virtualisation
- Apache/Tomcat in different VM's



Clinical data for research

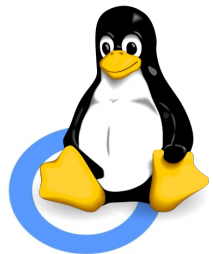
- ◆ Different data sources
 - ◆ laboratories
 - ◆ bedside data
 - ◆ different departments
 - ◆ multicentric/international research
- ◆ Strict regulations
- ◆ Complex research protocols
- ◆ Monitoring of research progress
- ◆ Long term archiving of data





Simple electronic forms (.xls, .mdb, ..)

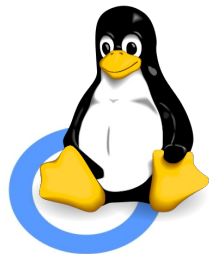
- + easy to develop
- + low price
- not practical for complex protocols
- bad for big research teams, multicentric studies
- without tracking changes, not compliant with regulations
- not flexible enough, lack of support for different data formats
- archiving not complete (signatures, modifications tracking)



„Tailored“ - custom made applications

- + web/desktop applications
- + everything can be customised

- time consuming
- \$\$\$ - price too high for small research teams
- bad reusability
- public clouds not an option (privacy, regulations)
- proprietary solutions (expensive, lock-in ...)



Essay

Could an Open-Source Clinical Trial Data-Management System Be What We Have All Been Looking For?

Greg W. Fegan, Trudie A. Lang*

Difficulties in Meeting the Demands of Regulators and Guidelines

In Europe, it is a legal requirement to conduct clinical trials in accordance with the International Conference on Harmonisation's guidelines on good clinical practice (see <http://www.ich.org/>). A recent editorial reported that this directive has led to a decline in the number of trials being conducted by independent academic groups [1]. One

Box 1. An Introduction to Open-Source Software: Definitions and Required Reading

1. Ten Things You Didn't Know about Open Source (<http://www.tectonic.co.za/view.php?id=1465>)
2. Definition of Open-Source Software
 - a. Free redistribution
 - b. Source code
 - c. Derived works

absence of guidance from regulatory agencies such as the European Medicines Agency and United States Food and Drug Administration about how to evaluate the many competing systems available, and indeed what the actual requirements are for trials where the data will be needed for a regulatory submission. This is particularly important with respect to trials evaluating products for neglected diseases, which are often carried out



Vanderbilt University

OpenClinica, LLC

Standard/LTS

Community/Enterprise

Available only for non-commercial research purposes

Lesser General Public License

Apache/PHP/MySql

Tomcat/J2EE/Postgres or Oracle

Certifikate FDA 21 part 11

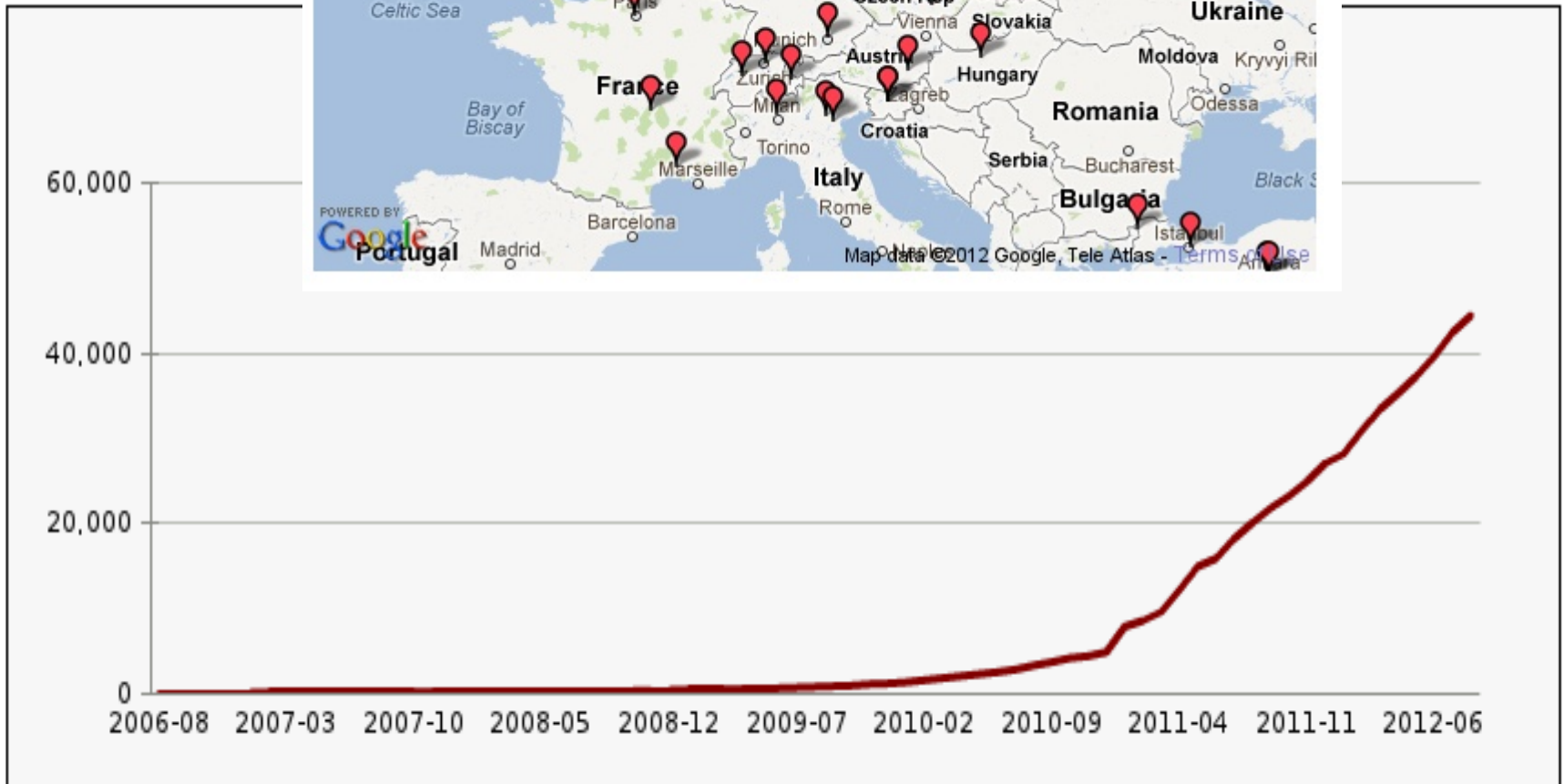
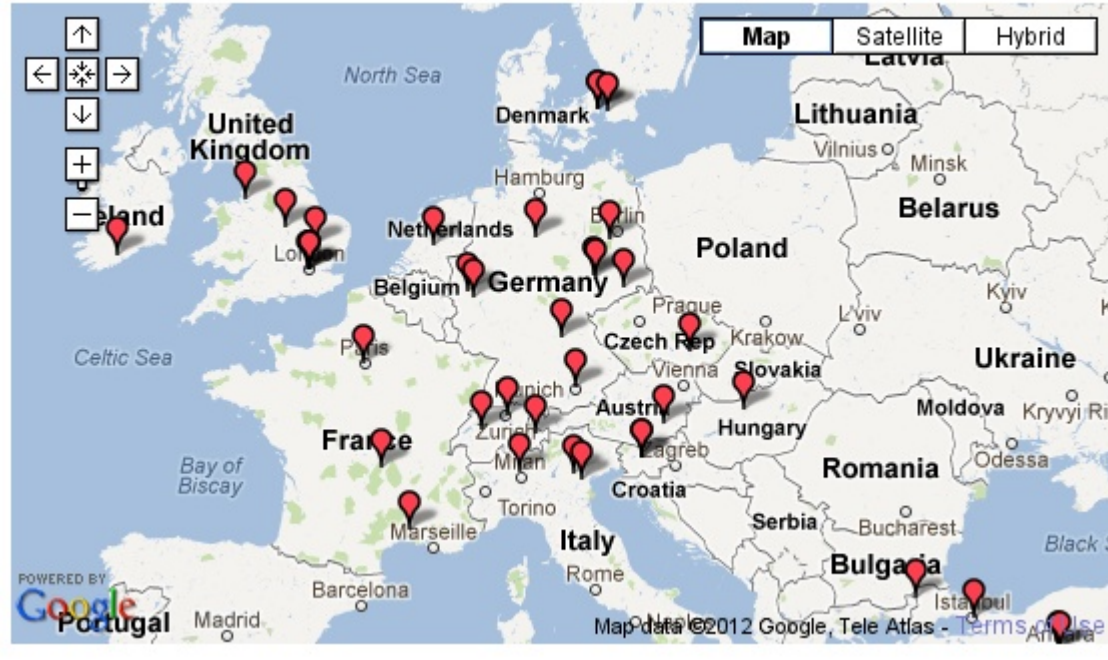
The screenshot shows the REDCap login interface. At the top left is the REDCap logo. Below it is a "Log In" link. A banner image features a hospital building and medical professionals, with text for "univerzitetni klinični center ljubljana" and "University Medical Centre Ljubljana". Below the banner, a message reads: "Please log in with your user name and password. If you are having trouble logging in, please contact [Iztok Štoll](#)." At the bottom, there are input fields for "Username:" and "Password:", a "Log In" button, and a "Forgot your password?" link.

The screenshot shows the OpenClinica login interface. At the top center is the OpenClinica logo with the tagline "Open Source for Clinical Research". Below the logo is a "Login" section with input fields for "User Name" and "Password", and a "Login" button with a "Forgot Password?" link. To the right is an "About" section with two bullet points: "▶ Testni raziskovalni strožnik Kliničnega oddelka za endokrinologijo, diabetes in bolezni presnove / UKC Ljubljana" and "▶ Klinični oddelk za endokrinologijo, diabetes in bolezni presnove". At the bottom, there is a footer with "OpenClinica Portal Help Contact", copyright information for Akaza Research LLC, and "Version: 3.0.1 Unsupported Community Edition" next to the Akaza Research logo.

Trends

Map of REDCap Consortium Partners

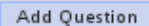
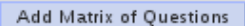
[View fullscreen map](#)





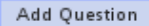
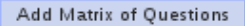
CRF (clinical research form) editor






Current instrument: **Prvi vprasalnik**

Preview instrument

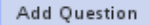

  **Demografski podatki**






 

     Variable: first_name

Ime

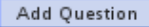
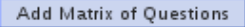
** must provide value*







 

     Variable: last_name

Priimek

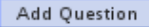

** must provide value*

      Variable: sex

Spol Ženski Moški

** must provide value* reset

Data Collection

Identifikacijska številka 1 [redacted]

Data Collection Instruments:

- Demographics
- Sprejem
- Anamneza Status
- Poseg In Zaključec
- Angiografija

Lock all forms

Applications

- Calendar
- Data Export Tool
- Data Import Tool
- Data Comparison Tool
- Logging
- File Repository
- User Rights
- Record Locking Customization
- E-signature and Locking Mgmt
- Graphical Data View & Stats
- Data Quality
- API
- Report Builder

Help & Information

- Help & FAQ
- Video Tutorials
- Suggest a New Feature



Univerzitetni klinični center Ljubljana
Klinični oddelek za endokrinologijo, diabetes in bolezni presnove

Angiografije diabetes 2012/1

Demographics

Share this instrument

VIDEO: Basic data entry (16 min)

Download PDF of - select PDF download option -

Editing existing Identifikacijska številka 1 [redacted]

Identifikacijska številka 1
(To rename this record, modify the value immediately below.)

Identifikacijska številka

Splošni podatki

Ime

Priimek

Datum rojstva Y-M-D

Spol

Form Status

Complete?

Lock this record for this form?

If locked, no user will be able to edit this record on this form until someone with Lock/Unlock privileges unlocks it.

Lock

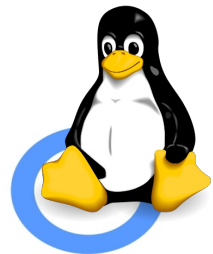
Additional features

Redcap:

- Tablet computers
- Data Transfer Service (DTS) – push or pull data
- Randomisation module
- Very friendly user interface
- Surveys

Openclinica:

- Very good REST/SOAP infrastructure / easy customisation
- Discrepancy management (GUI entry and WS)
- CDISC ODM



Trenutna telesna višina v cm : [Refresh Plot](#)

Katere vrste športa ali rekreacije se navadno udeležujete ? : [Refresh Plot](#)

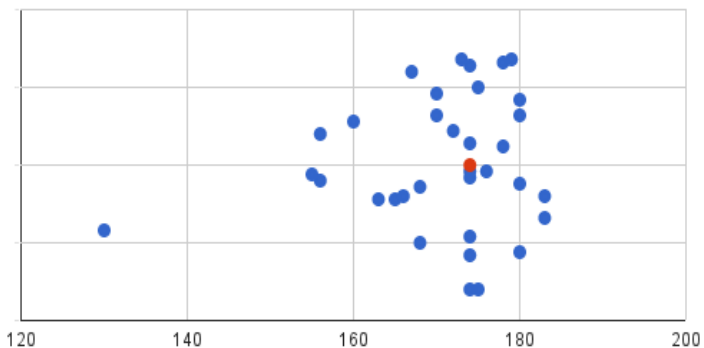
View as Bar Chart ▾

Total (N)	Missing	Unique	Min	Max	Mean	StDev	Percentile					
							.05	.10	.25	.50 Median	.75	.90
34	1 (2.9%)	19	130.00	183.00	170.71	10.39	142.50	156.00	166.50	174.00	178.00	180.00

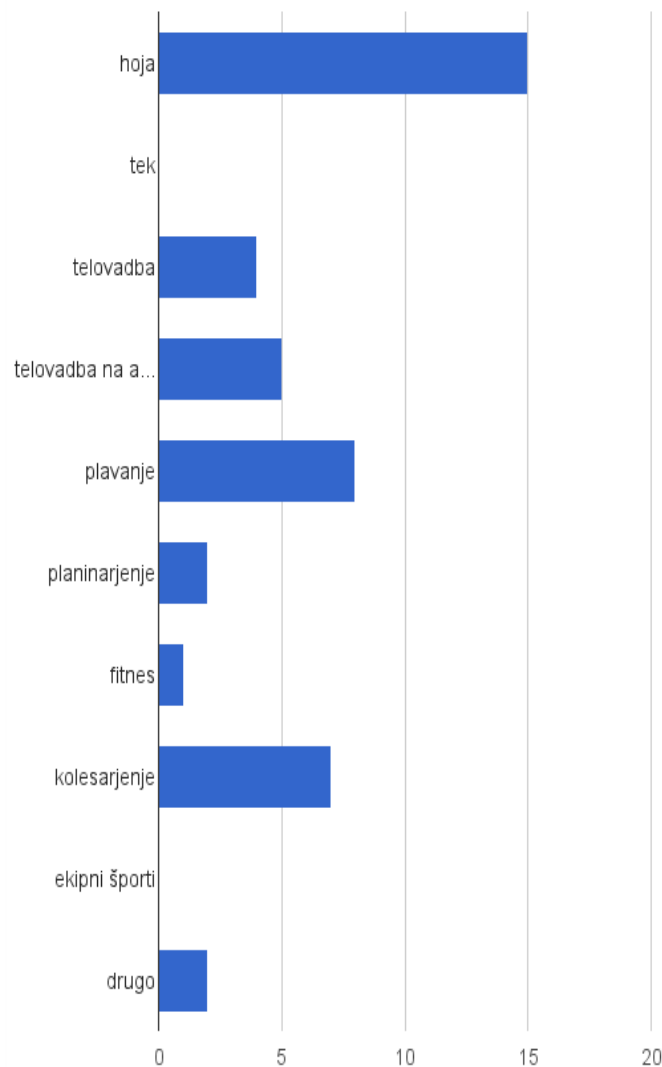
Total (N)	Missing	Unique
20	15 (42.9%)	1

Lowest values: 130, 155, 156, 156, 160

Highest values: 180, 180, 180, 183, 183



Counts/frequency: hoja (15, 75%), tek (0, 0%), telovadba (4, 20%), telovadba na aparatih doma (Orbitrek, sobno kolo) (5, 25%), plavanje (8, 40%), planinarjenje (2, 10%), fitnes (1, 5%), kolesarjenje (7, 35%), ekipni športi (0, 0%), drugo (2, 10%)

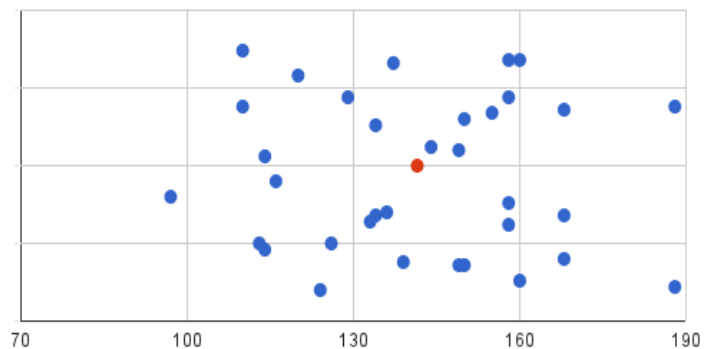


Trenutna telesna teža v kg: [Refresh Plot](#)


















Total (N)	Missing	Unique	Min	Max	Mean	StDev	Percentile						
							.05	.10	.25	.50 Median	.75	.90	
34	1 (2.9%)	22	97.00	188.00	141.62	22.69	103.50	111.50	122.00	141.50	158.00	168.00	1

Lowest values: 97, 110, 110, 113, 114

Highest values: 168, 168, 168, 188, 188



Data export

		Download Syntax & Data
	Microsoft Excel You may download the survey results in CSV (comma-separated) format, which can be opened in Excel. You have the choice of downloading the data either with the full headers and answer labels or just with the answer codes (i.e. raw data). <i>NOTE: If you are using a version of Microsoft Excel prior to Excel 2007, due to limitations the data will only be read to 255 columns when opened.</i>	 EXCEL CSV Labels  EXCEL CSV Raw <input type="checkbox"/> Send file?
	SPSS Statistical Analysis Software Instructions: Download and save all 3 files on the right to a common location. First, double-click on the Pathway Mapper (.bat) file, which will run quickly and invisibly. (If you are not using a Windows operating system, such as Mac or Linux, please see the <i>Additional Instructions</i> .) Now double-click on the *.sps file, which will open SPSS. When the file is loaded and displayed, choose Run-->All from the top menu options. This action will launch the script that will automatically read in all data and manipulate data fields with labels, option values, etc. Additional instructions	 SPSS  DATA CSV  Pathway Mapper <input type="checkbox"/> Send file?
	SAS Statistical Software Instructions: Download and save all 3 files on the right to a common location. First, double-click on the Pathway Mapper (.bat) file, which will run quickly and invisibly. (If you are not using a Windows operating system, such as Mac or Linux, please see the <i>Additional Instructions</i> .) Now double-click on the *.sas file, which will open SAS. When the file is loaded and displayed, choose Run (or Run-->Submit) from the top menu options. This action will launch the script that will automatically read in all data and manipulate data fields with labels, option values, etc. Additional instructions	 SAS  DATA CSV  Pathway Mapper <input type="checkbox"/> Send file?
	R Statistical Software Instructions: Use command <code>read.csv('filename')</code> to read in data file.	 R  DATA CSV <input type="checkbox"/> Send file?
	STATA Analysis and Statistical Software Instructions: Download both files to common location and double-click on *.do file. This action will launch the script that will automatically read in all data and manipulate data fields with labels, option values, etc.	 STATA  DATA CSV <input type="checkbox"/> Send file?










Alerts & Messages

Instructions

Info

Icon Key

Statutes

-  Not Started
-  Scheduled
-  Data Entry Started
-  Stopped
-  Skipped
-  Completed
-  Signed
-  Locked
-  Invalid

Actions

-  View
-  Edit
-  Remove
-  Restore
-  Reassign
-  Sign

[View All Icons](#)

View Subjects in NSB 2011 ?

Study Subject ID	Subject Status	Site ID	OID	Sex	Secondary ID	Prvi pregled	Kontrolni obisk	Po porodu	Actions
									Apply Filter Clear Filter
24042023	available	Poliklinika - NSB 2011	SS_24042023	f			x2		
24042126	available	Poliklinika - NSB 2011	SS_24042126	f					
24042042	available	Poliklinika - NSB 2011	SS_24042042	f					
24042181	available	Poliklinika - NSB 2011	SS_24042181	f					
24042045	available	Poliklinika - NSB 2011	SS_24042045	f			x4		
24042123	available	Poliklinika - NSB 2011	SS_24042123	f			x3		
24042146	available	Poliklinika - NSB 2011	SS_24042146	f					
24042099	available	Poliklinika - NSB 2011	SS_24042099	f			x2		
24042190	available	Poliklinika - NSB 2011	SS_24042190	f					
24042108	available	Poliklinika - NSB 2011	SS_24042108	f			x2		
24042137	available	Poliklinika - NSB 2011	SS_24042137	f			x3		
24042087	available	Poliklinika - NSB 2011	SS_24042087	f			x4		
24042140	available	Poliklinika - NSB 2011	SS_24042140	f					
24042166	available	Poliklinika - NSB 2011	SS_24042166	f			x2		
24042081	available	Poliklinika - NSB 2011	SS_24042081	f			x3		

Subject: 24042023 X

Event: Prvi pregled

Status: data entry started

- View/Enter Data
- Edit
- Remove

Conclusions

+

- Excellent stability
- Very good support/very active community
- Low cost of ownership

-

- ?

"EUropean Best Information through Regional Outcomes in Diabetes" (EUBIROD)



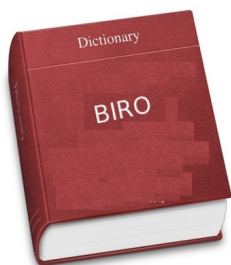
Public health project in the field of diabetes sponsored by the European Union

Goal: to create extensive report on several regions in Europe in diabetes care

EUBIROD



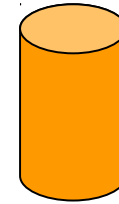
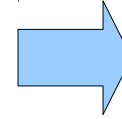
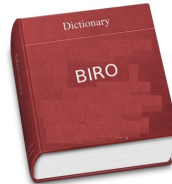
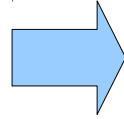
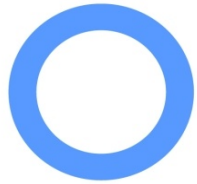
BIRO Core EU Dataset



N=48

1. **ID Patient**
2. ID Centre
3. **Type of Diabetes**
4. **Sex**
5. **Date of Birth**
6. **Date of Diagnosis**
7. **Episode Date**
8. Smoking Status
9. N.Cigarettes (x day)
10. Alcohol Intake (g/x day)
11. Weight
12. Height
13. BMI
14. Systolic Blood Pressure
15. Diastolic Blood Pressure
16. HbA1c
17. Creatinine
18. Microalbumin
19. Total Cholesterol
20. HDL
21. Tryglicerides
22. Eye Examination
23. Retinopathy Status
24. Maculopathy Status
25. Foot Examination
26. Foot Pulses
27. Foot vibration
28. End Stage Renal Failure
29. Renal Dyalysis
30. Renal Transplant
31. Stroke
32. Foot Ulceration
33. Acute Myocardial Infarction
34. Laser
35. Hypertension
36. Blindness
37. Amputation
38. Antihypertensive Medication
39. Hypoglycemic Drug Therapy
40. Oral Drug Therapy
41. Pump Therapy
42. Nasal Therapy
43. Average Injections (x day)
44. Self monitoring
45. Diabetes Specific Education
46. Lipid Lowering Therapy
47. Anti-platelet Therapy
48. Patient enrollment in DMP for diabetes

1. step: Discard Heterogeneity from primary sources



Diabetes
clinic/region/country

Mapping to
standard

Standardized
database



BIROBox

Help

B.I.R.
Best Information through Regional Outcomes

BIROBox
Setup

BIRO Database
Database Engine

Local Report
Statistical Engine

Global Report
Central Engine

Data Transmission
Communication Software

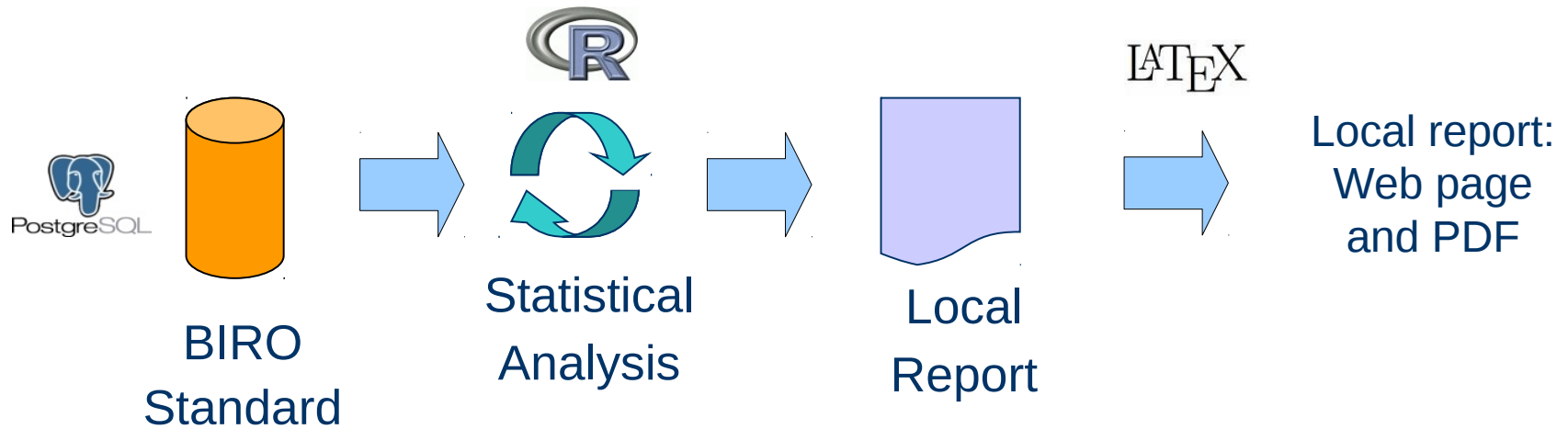
Fields mapping configuration

Configure mapping between BIRO fields and local fields

BIRO field	BIRO field name: DOB
Date of Birth	BIRO field code: BIRO005
Date of Diagnosis	BIRO field description: date or birth
Patient ID	Lower threshold: 1900-01-01
Sex	<input checked="" type="checkbox"/> Extract from local database
Sub-Data Source ID	Local field name
Type of Diabetes	dbirth
Alcohol Intake	Select the input date format
Alcohol status	yyyy-MM-dd
Amputation	
Anti Platelet Therapy	
Average Injections	
Biguanide therapy	
Blindness	
BMI	
Cigarettes per day	
Creatinine	
Diabetes Specific Educ...	
Diastolic blood-pressure	
End Stage Renal Failure	
Episode Date	
Eye Examination	
Foot Examination	
Foot Pulses	
Foot Sensation	
Foot Ulcer	
Glinide therapy	
Glitazone therapy	
Glucosidase inhibitor t...	
HbA1c	
HDL	
Height	
Hypertension	
Hypertensive Medication	
Hypoglycaemic Drug Th...	
Laser	

Previous Finish

2. step: Local statistical analysis



The screenshot shows the BIROBox application window. On the left is a dark blue sidebar with the BIRO logo and a navigation menu containing: BIROBox (Setup), BIRO Database (Database Engine), Local Report (Statistical Engine), Global Report (Central Engine), and Data Transmission (Communication Software). The main area is titled 'Statistical Engine Configuration' and contains the following fields:

- BIRO Database: dropdown menu with 'birox' selected and a 'Refresh' button.
- Centre ID: text input field with 'SI.1'.
- Centre name: text input field with '[European Union].[Slovenia]'.
- Start year: spinner box with '2012'.
- Duration (years): spinner box with '1'.
- Reference date: dropdown menu with '12-31'.
- Sub data sources: checkbox labeled 'enable sub data sources reporting' (unchecked).

At the bottom of the window are two buttons: 'Run Statistical Engine' and 'Browse Results'.

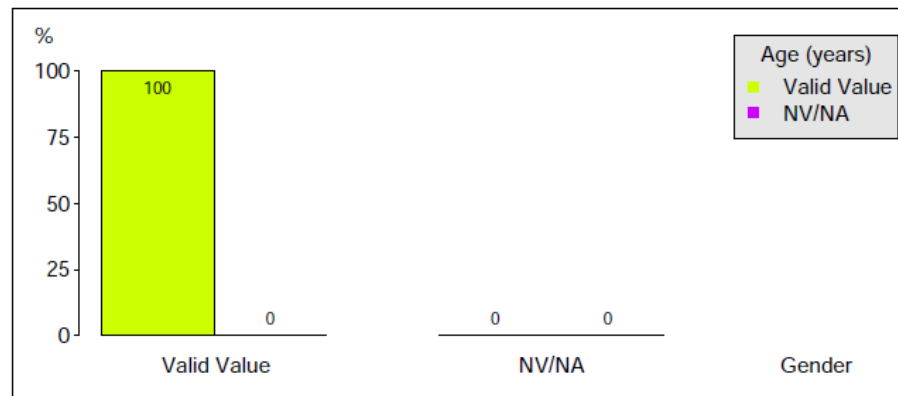
Contents

<u>1</u>	<u>Demographic characteristics</u>	6
1.1	Basic demographics	7
1.1	Age (Classes)	8
<u>2</u>	<u>Clinical characteristics</u>	11
2.1	Diabetes Status	12
2.1.1	Type of diabetes	13
2.1.2	Duration of diabetes (Classes)	17
2.2	Risk Factors	40
2.2.1.1.	Weight (last episode in 12 months)	41
2.2.1.2.	BMI (last episode in 12 months)	98
2.2.2	Lifestyle	149
2.2.2.1	Smoking status (last episode in 12 months)	150
2.2.3.	Clinical measurements	180
2.2.3.1.	Systolic BP (last episode in 12 months)	181
2.2.3.2.	Diastolic BP (last episode in 12 months)	237
2.2.3.3	Total cholesterol (last episode in 12 months)	293
2.2.3.4.	HDL-cholesterol (last episode in 12 months)	326
2.2.3.5	Creatinine (last episode in 12 months)	359
2.2.3.6.	HbA1c (last episode in 12 months)	416
2.3	Diabetes complications	472
2.3.1.	Retinopathy (first episode in 12 months)	473
2.3.2.	End stage renal failure (first episode in 12 months)	477
2.3.3.	Foot ulcer (first episode in 12 months)	481
2.3.4.	Amputation (first episode in 12 months)	485
2.3.5.	Stroke (first episode in 12 months)	489
2.3.6.	Myocardial infarction (first episode in 12 months)	493
2.3.7.	Hypertension (first episode in 12 months)	496
<u>3</u>	<u>Health System</u>	500
3.1.	Structure (provider level)	501
3.1.1	Type of Provider	502
3.1.2	Average diabetes population	503
3.2.	Structural quality	515
3.2.1	Hospital beds per 100,000 population	516
3.2.2	Physicians employed per 100,000 population	517
3.3.	Processes (individual level)	518
3.3.1.	Foot examination	519
3.3.1.1	Done (last episode in 12 months)	520
3.3.2.	Eye examination	539
3.3.2.1	Done (last episode in 12 months)	540
3.3.3	Measurements examination	559
3.3.3.1	BP (last episode in 12 months)	560
3.3.3.2	Lipids	579

Age	Valid Value (%)	NV/NA (%)	N (%)
Valid Value	118156 (100.0)	0(0.0)	118156 (100.0)
NV/NA	0 (0.0)	0(0.0)	0 (0.0)
TOTAL	118156(100.0)	0(0.0)	118156 (100.0)

Table 1.1.1.1 : Age (by Gender)

CMH Chi-Square
Value Too many cells have 0 obs

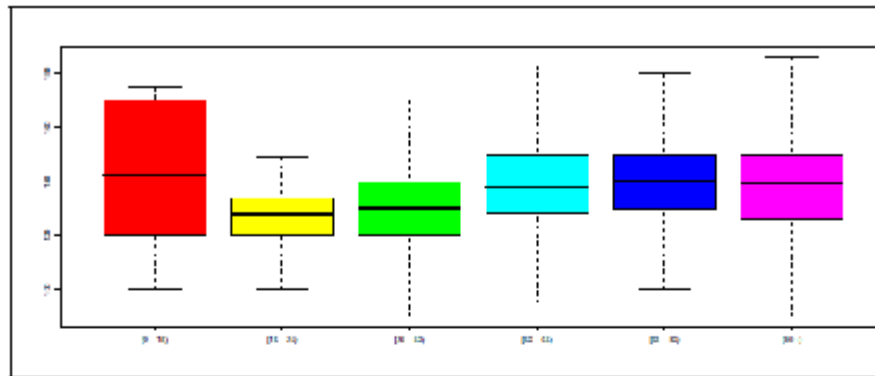


Barplot 1.1.1.1 - Age (by Gender)

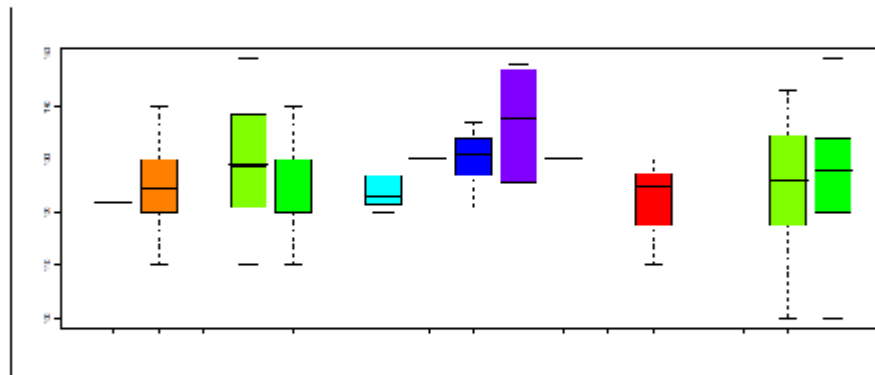
Age	Male (%)	Female (%)	N (%)
[0 - 15)	482 (0.8)	490(0.9)	972 (0.8)
[15 - 25)	1104 (1.8)	1009(1.8)	2113 (1.8)
[25 - 50)	8867 (14.3)	6613(11.8)	15480 (13.1)
[50 - 65)	22623 (36.5)	16056(28.6)	38679 (32.7)
[65 - 80)	23731 (38.2)	23834(42.5)	47565 (40.3)
[80 +)	5249 (8.5)	8097(14.4)	13346 (11.3)
TOTAL	62056(52.5)	56099(47.5)	118155 (100.0)

Table 1.1.1.2 : Age (by Gender)

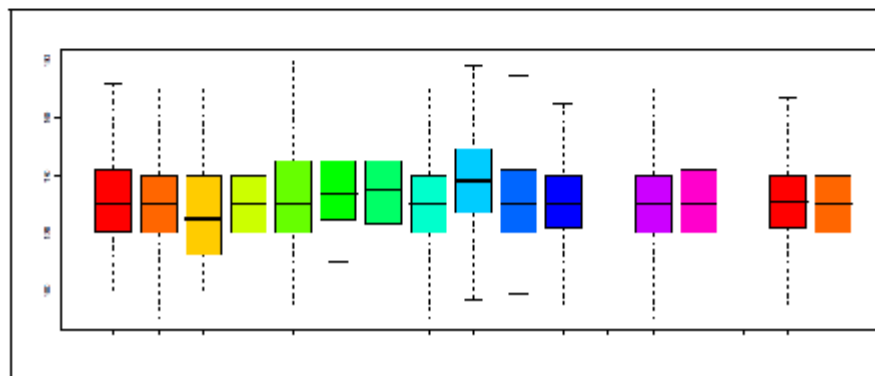
CMH Chi-Square p.value df
Value 1759.6126 0 5



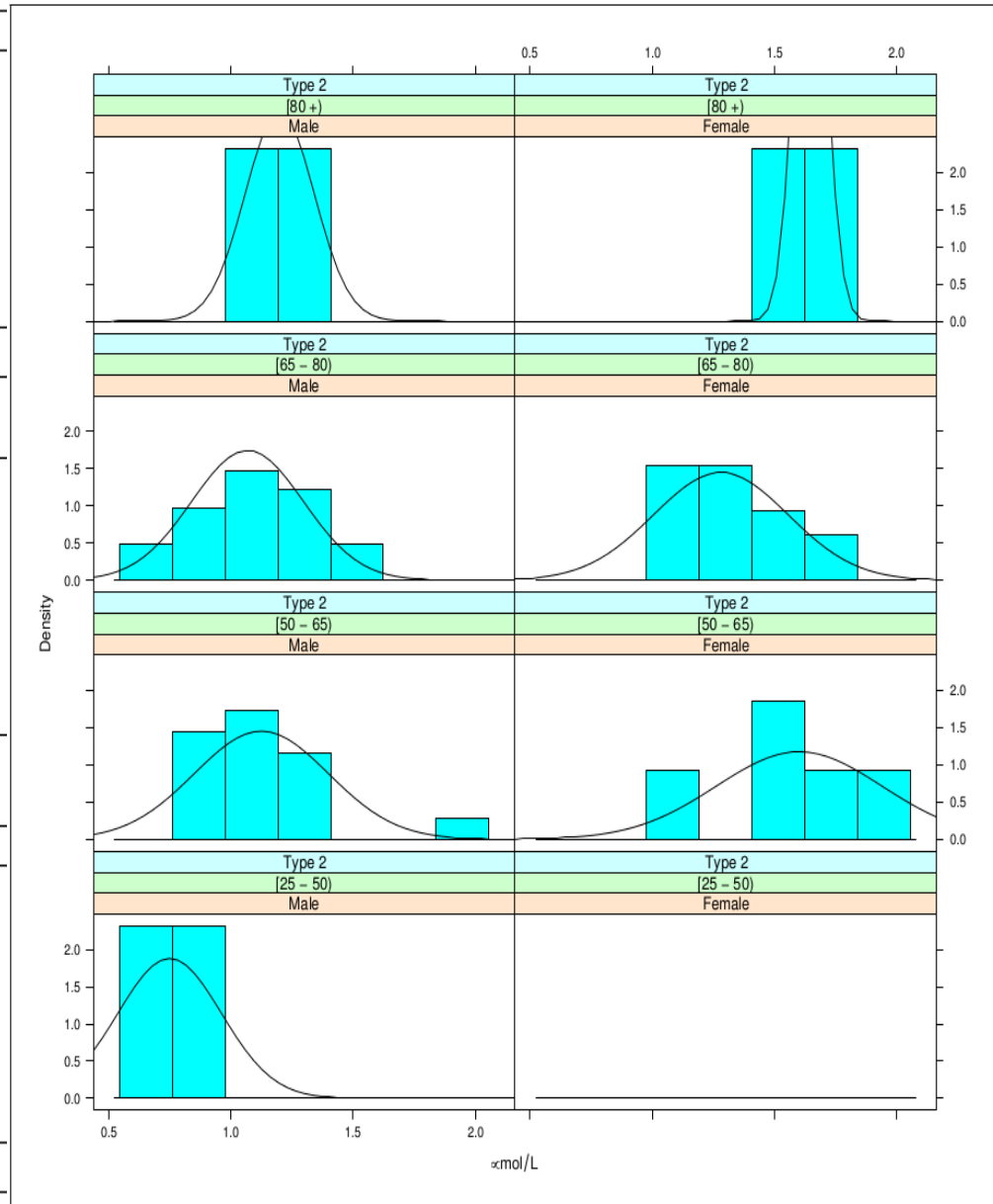
Boxplot: 2.2.3.1.30 - SBP (by Age, Type of Diabetes = Type 2)



Boxplot: 2.2.3.1.31 - SBP (by Data sources, Age = [15 - 25], Type of Diabetes = Type 2)



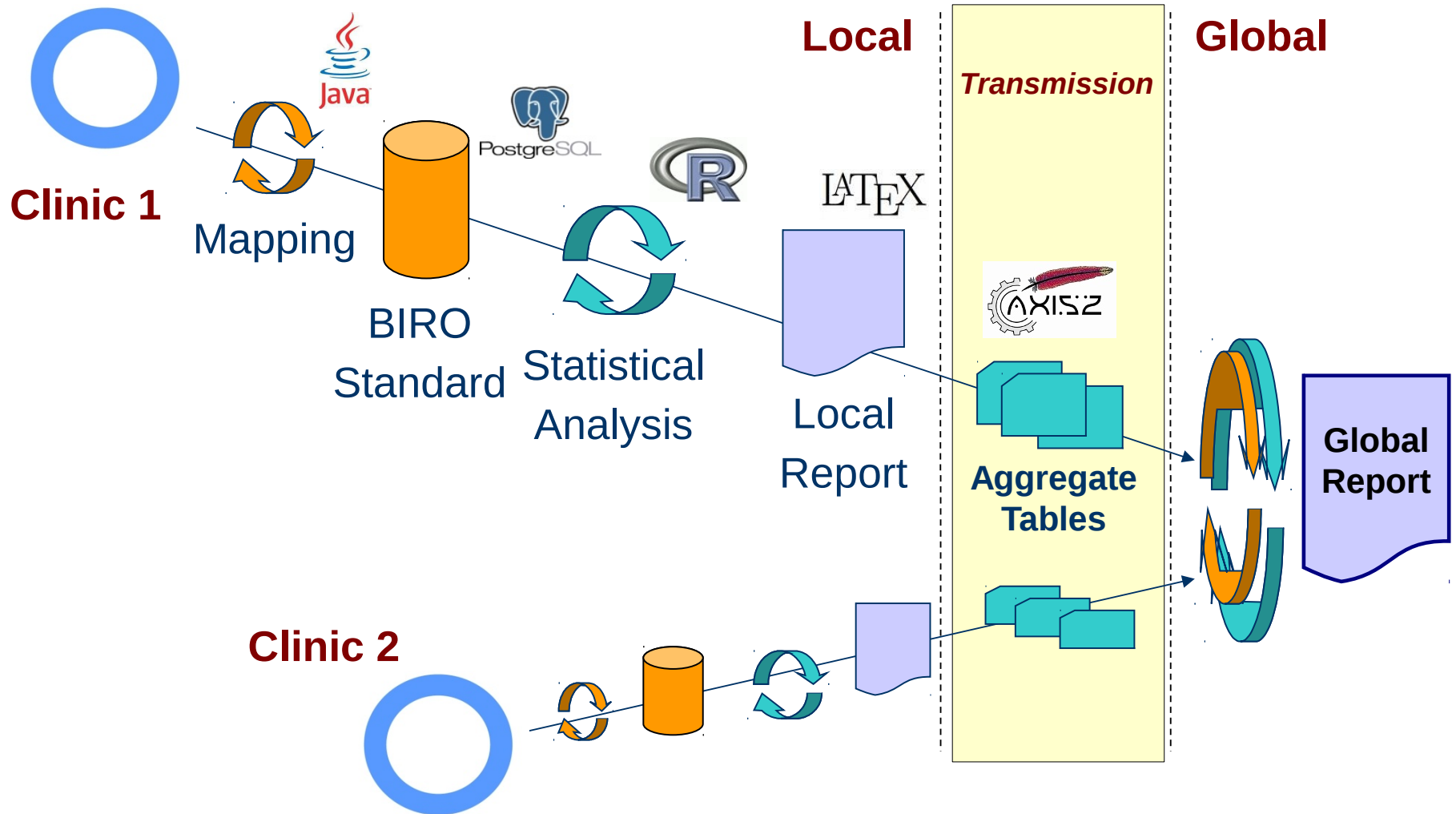
Boxplot: 2.2.3.1.32 - SBP (by Data sources, Age = [25 - 60], Type of Diabetes = Type 2)



Trellis density plot: 2.2.3.4.2 - HDL * Gender * Age (Type of Diabetes = Type 2)

The complete BIRO model

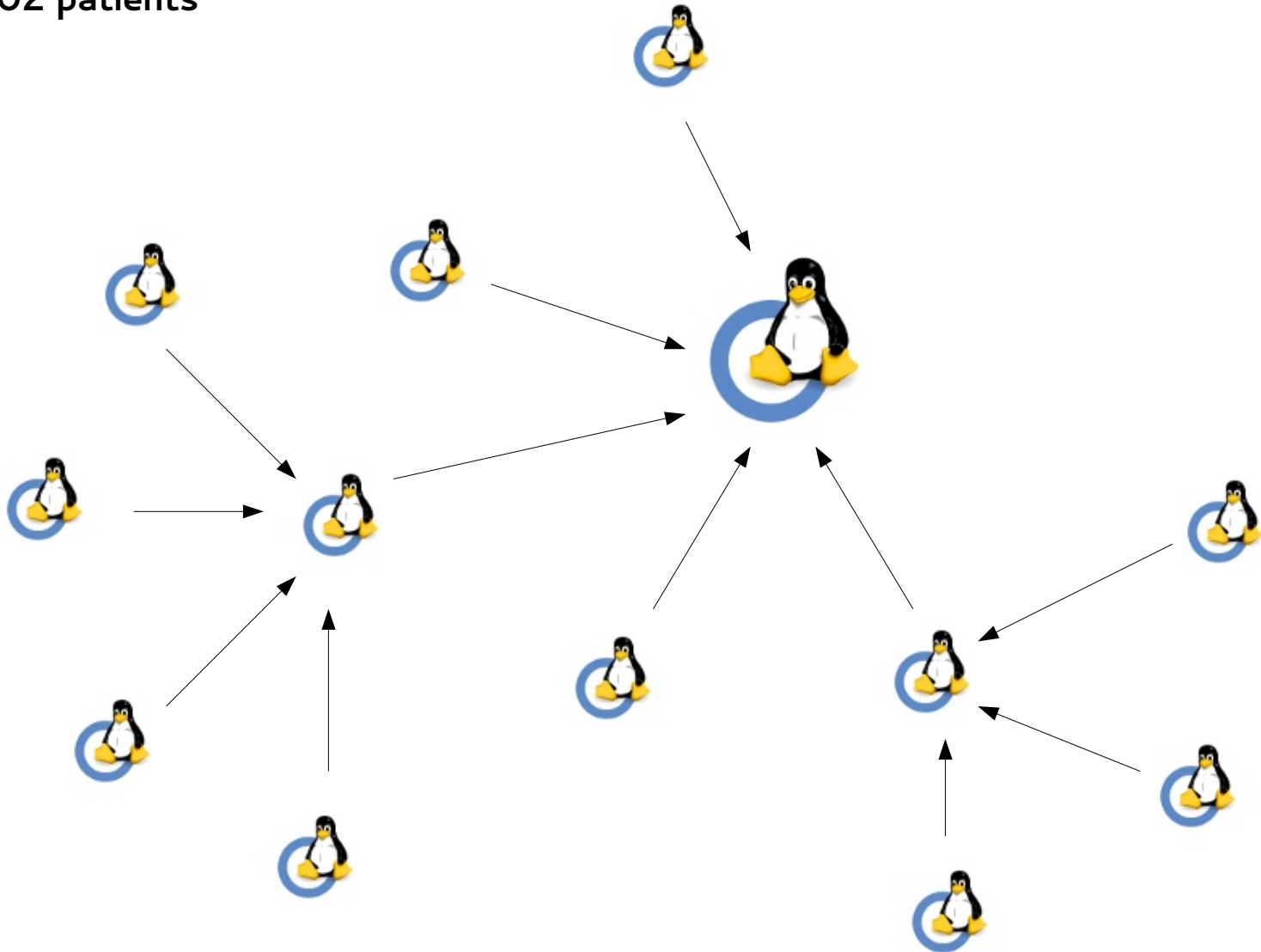
www.biro-project.eu



Report 2010:

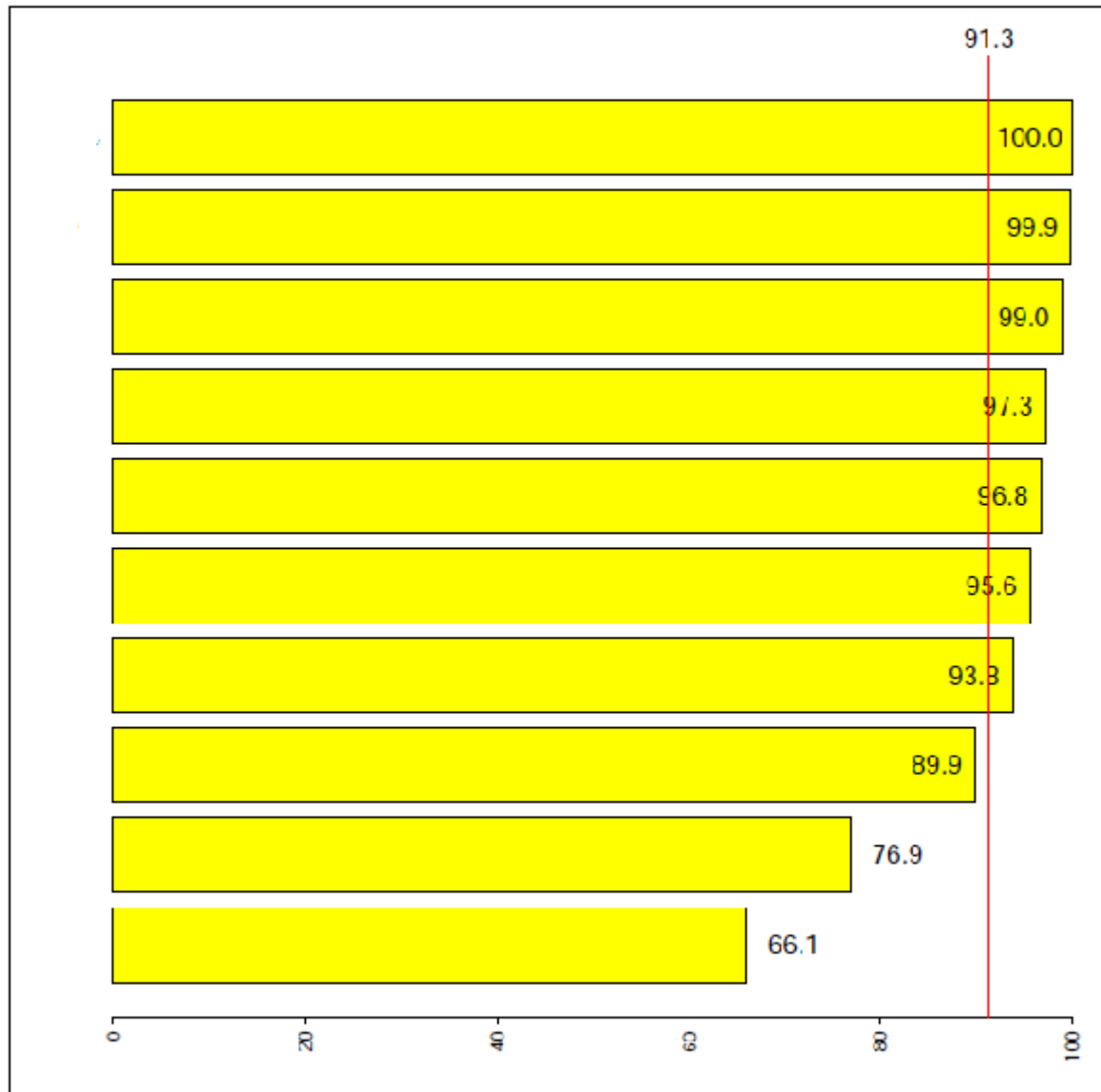
Centres from 19 european countries

N=199,902 patients



5.2.4 % of subjects receiving at least one foot examination within the last 12 months

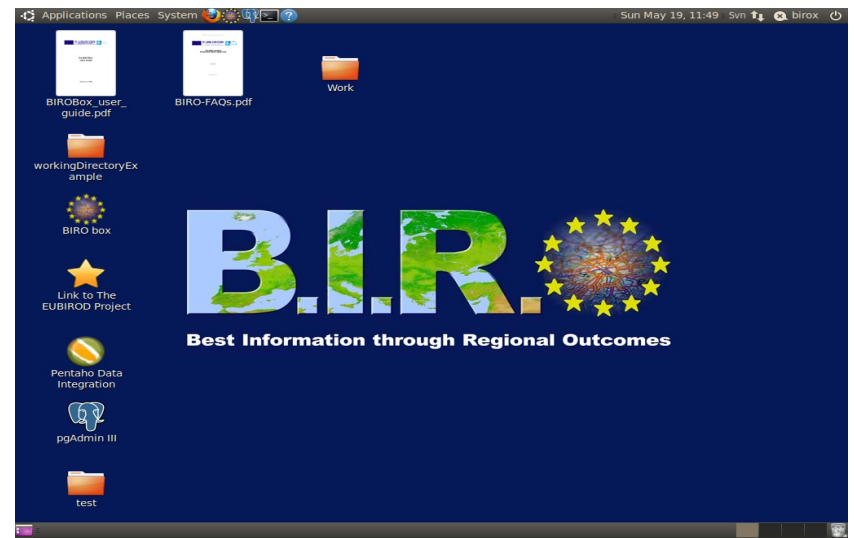
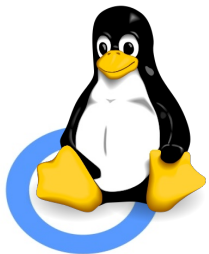
Type of Diabetes = Type 2



Barplots: 5.2.4.239 - Adjusted Rates 5.2.4 % of subjects receiving at least one foot examination within the last 12 months

BIROX Linux

- Solves problem of software distribution and upgrades
- VirtualBox Linux appliance – based on Ubuntu
- Upgrades of software through repository/system update
- Includes:
 - BIROBox – Java desktop Client
 - Preconfigured postgres database, R, LaTeX, Java
 - Documentation
 - Some additional tools



Collaboration



Mediawiki: distribution of EHR data model proposal

www.EndoDiabosi

dotCMS (opensource java CMS):

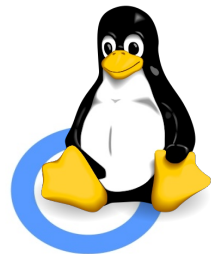
- distribution of clinical guidelines
- education materials
- events
- patient information site



Alfresco: paper forms repository



Moodle: education



Open source development model

- easier debugging
- feature additions and customization for special needs
- transparency of patient data manipulation
- free as in beer (otherwise often very expensive and unaffordable to small departments)



Questions ...

