EUBIROD

EUropean Best Information through
Regional Outcomes in Diabetes

Cross-border Flow of Health Information: is “Privacy by Design” sufficient to obtain complete and accurate data for Public Health in Europe?

The case of BIRO/EUBIROD Diabetes Registers

Concetta Tania Di Iorio
Serectrix snc
on behalf of the EUBIROD Consortium

3rd European Public Health Conference
Amsterdam 12th November 2010

Directorate-General for
Health & Consumers

Executive
Agency for
Health and
Consumers

The BIRO Project

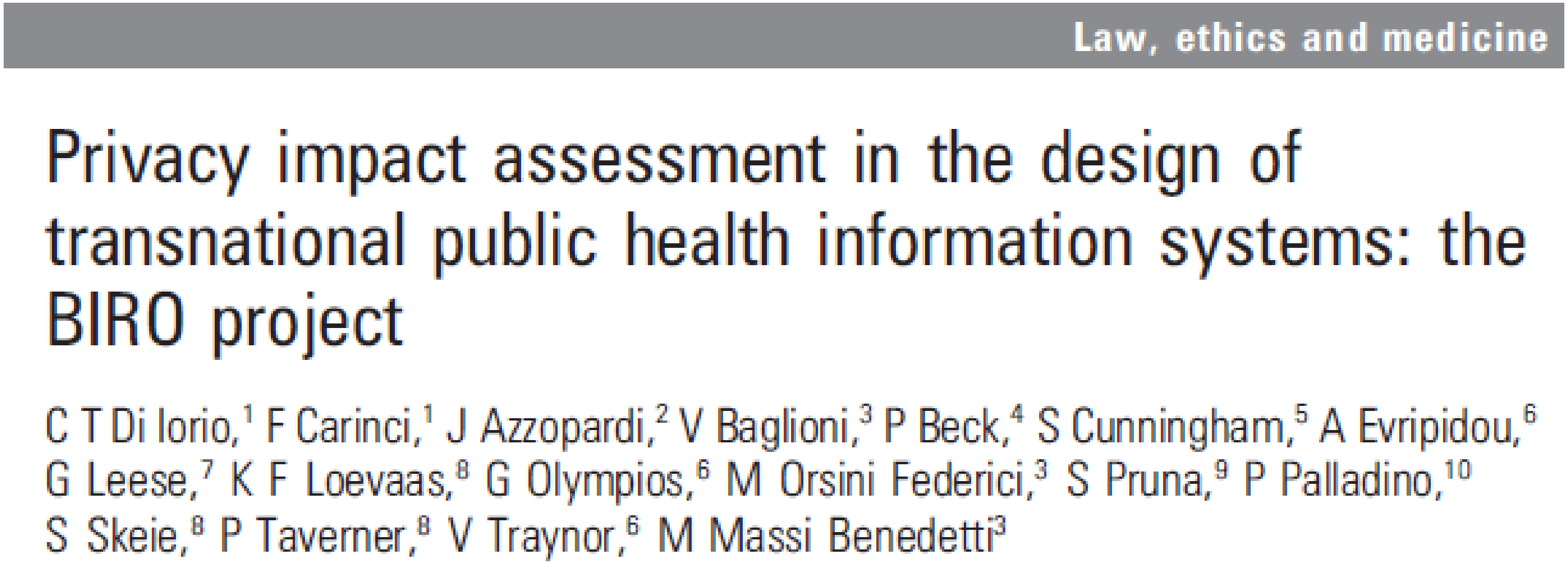
- **General Aim:** ***to build a common European infrastructure for the routine production of quality and outcome indicators through the standardized and secure exchange of information across regional diabetes registers***
- **Specific Aim**: to implement the concept of **"Privacy by Design":**
    - **privacy issues and concerns identified from the early design stage**
    - **mitigation strategies directly implemented in the system architecture**

Privacy Impact Assessment

- The BIRO Consortium conceived and applied **a novel method of Privacy Impact Assessment (PIA)** to fulfil **“Privacy by Design”**
- **Selection of the best system architecture** in terms of:
  - **privacy protection**
  - **information content**
  - **technical complexity (feasibility)**

BIRO Infrastructure: “Privacy by Design”

DI IORIO CT et al, J Med Ethics. 2009 Dec;35(12):753-61.

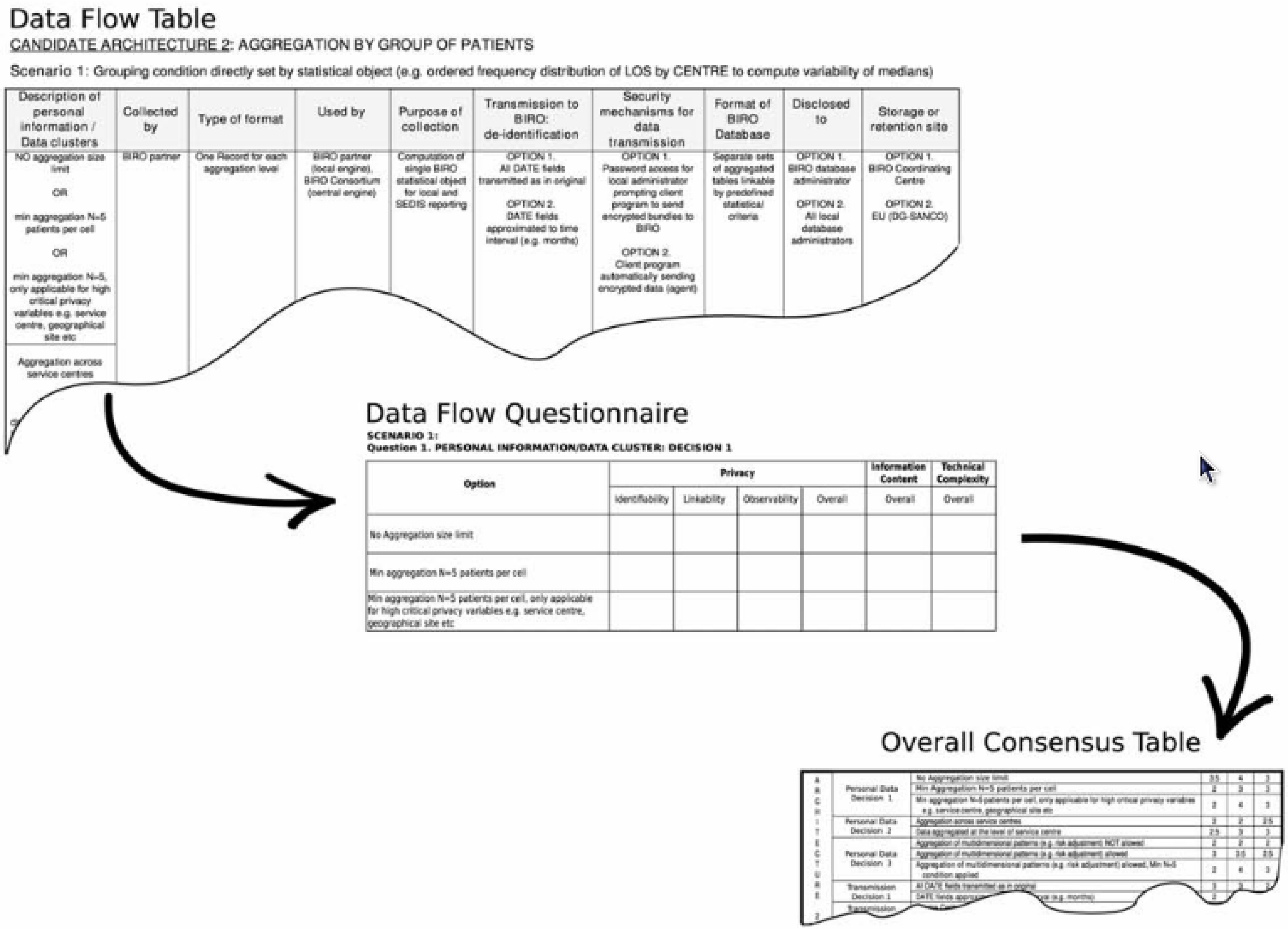


Law, ethics and medicine

Privacy impact assessment in the design of transnational public health information systems: the BIRO project

C T Di Iorio,[1] F Carinci,[1] J Azzopardi,[2] V Baglioni,[3] P Beck,[4] S Cunningham,[5] A Evripidou,[6] G Leese,[7] K F Loevaas,[8] G Olympios,[6] M Orsini Federici,[3] S Pruna,[9] P Palladino,[10] S Skeie,[8] P Taverner,[8] V Traynor,[6] M Massi Benedetti[3]

Procedure



Data Flow Table

CANDIDATE ARCHITECTURE 2: AGGREGATION BY GROUP OF PATIENTS

Scenario 1: Grouping condition directly set by statistical object (e.g. ordered frequency distribution of LOS by CENTRE to compute variability of medians)

Description of personal information / Data clusters	Collected by	Type of format	Used by	Purpose of collection	Transmission to BIRO: de-identification	Security mechanisms for data transmission	Format of BIRO Database	Disclosed to	Storage or retention site
NO aggregation size limit OR min aggregation N=5 patients per cell OR min aggregation N=5, only applicable for high critical privacy variables e.g. service centre, geographical site etc	BIRO partner	One Record for each aggregation level	BIRO partner (local engine), BIRO Consortium (central engine)	Computation of single BIRO statistical object for local and SEDIS reporting	OPTION 1. All DATE fields transmitted as in original OPTION 2. DATE fields approximated to time interval (e.g. months)	OPTION 1. Password access for local administrator prompting client program to send encrypted bundles to BIRO OPTION 2. Client program automatically sending encrypted data (agent)	Separate sets of aggregated tables linkable by predefined statistical criteria	OPTION 1. BIRO database administrator OPTION 2. All local database administrators	OPTION 1. BIRO Coordinating Centre OPTION 2. EU (DG-SANCO)
Aggregation across service centres									

Data Flow Questionnaire

SCENARIO 1:
Question 1. PERSONAL INFORMATION/DATA CLUSTER: DECISION 1

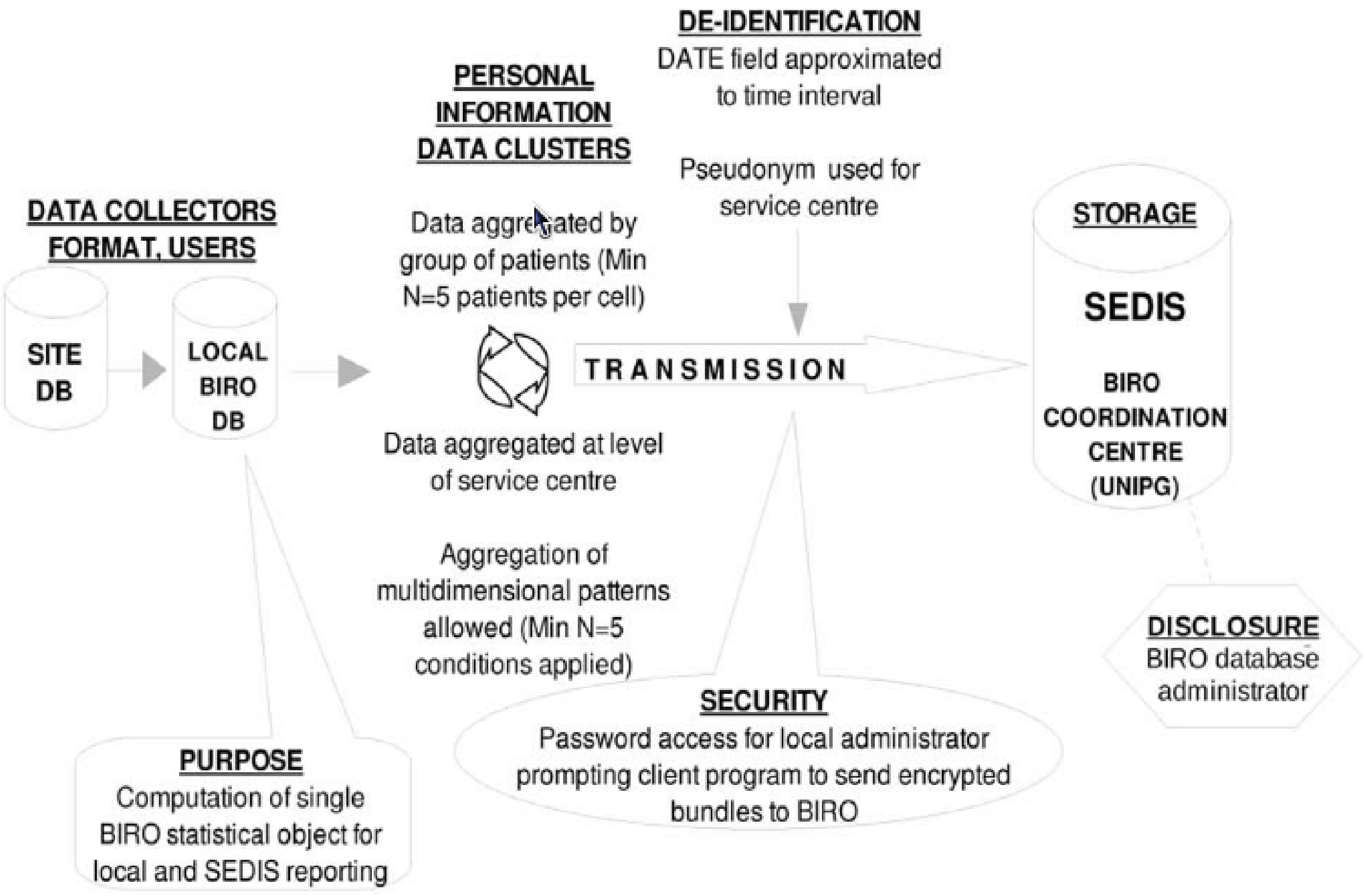
Option	Privacy				Information Content	Technical Complexity
	Identifiability	Linkability	Observability	Overall	Overall	Overall
No Aggregation size limit						
Min aggregation N=5 patients per cell						
Min aggregation N=5 patients per cell, only applicable for high critical privacy variables e.g. service centre, geographical site etc						

Overall Consensus Table

ARCHITECTURE 2	Personal Data Decision 1	No Aggregation size limit	3.5	4	3
		Min Aggregation N=5 patients per cell	2	3	3
		Min aggregation N=5 patients per cell, only applicable for high critical privacy variables e.g. service centre, geographical site etc	2	4	3
	Personal Data Decision 2	Aggregation across service centres	2	2	2.5
		Data aggregated at the level of service centre	2.5	3	3
	Personal Data Decision 3	Aggregation of multidimensional patterns (e.g. risk adjustment) NOT allowed	2	2	2
		Aggregation of multidimensional patterns (e.g. risk adjustment) allowed	3	3.5	2.5
		Aggregation of multidimensional patterns (e.g. risk adjustment) allowed, Min N=5 condition applied	2	4	3
	Transmission Decision 1	All DATE fields transmitted as in original	3	3	2
		DATE fields appro[illegible] (e.g. months)	2		
	Transmission				

Architecture of the BIRO System

Di Iorio CT et al., J Med Ethics. 2009 Dec;35(12):753-61.



DATA COLLECTORS
FORMAT, USERS

SITE DB

LOCAL BIRO DB

PERSONAL INFORMATION DATA CLUSTERS

Data aggregated by group of patients (Min N=5 patients per cell)

Data aggregated at level of service centre

Aggregation of multidimensional patterns allowed (Min N=5 conditions applied)

DE-IDENTIFICATION
DATE field approximated to time interval

Pseudonym used for service centre

TRANSMISSION

STORAGE

SEDIS

BIRO COORDINATION CENTRE (UNIPG)

DISCLOSURE
BIRO database administrator

PURPOSE
Computation of single BIRO statistical object for local and SEDIS reporting

SECURITY
Password access for local administrator prompting client program to send encrypted bundles to BIRO

Privacy Impact Assessment Report Conclusions

- The BIRO architecture fulfils privacy protection requirements by addressing and resolving **broad privacy concerns** from different angles**:**
    - ***individual's privacy + legal entities' privacy***
- The BIRO project attempts to reach the **best trade-off between the right to privacy and the right to better health care:**
    - *fully respectful of individual rights by exchanging only anonymous data*
    - *without jeopardizing information content for public health*
- **The BIRO Privacy Impact Assessment approach may represent a general methodology for the design of trans-border health information systems**

The EUBIROD Project

The **EUBIROD project (2008-2011)** aims:

- to implement a sustainable **European Diabetes Register** through the coordination of existing national/regional frameworks
- to systematically **use the BIRO technology in 20 European countries** to deliver **European Diabetes Reports** on a regular basis

The EUBIROD Privacy Impact Assessment

- **General Aim: to document the impact of the BIRO system in the broader / heterogeneous context of the EUBIROD Consortium**
- **Specific Aims:**
  - identification of **key elements of data protection**
  - **classification of key elements into factors/sub-factors**
  - creation of a **questionnaire** to collect information on data processing
  - **analysis of the variability of approaches** across Europe
  - development of an **IT platform to improve the management of privacy issues** in the management of disease registers
- The fulfillment of these activities allowed to ascertain:
  - **heterogeneity in the implementation of privacy principles/requirements**
  - **key areas of concern**

EUBIROD Privacy Impact Assessment Questionnaire

Includes N=11 sections - one for each **factor** identified.

Each section (factor) includes various questions **(sub-factors)**

FACTORS:

A1. Accountability of personal information

A2. Collection of Personal Information

A3. Consent

A4. Use of Personal Information

A5. Disclosure and Disposition of Personal Information

A6. Accuracy of Personal Information

A7. Safeguarding Personal Information

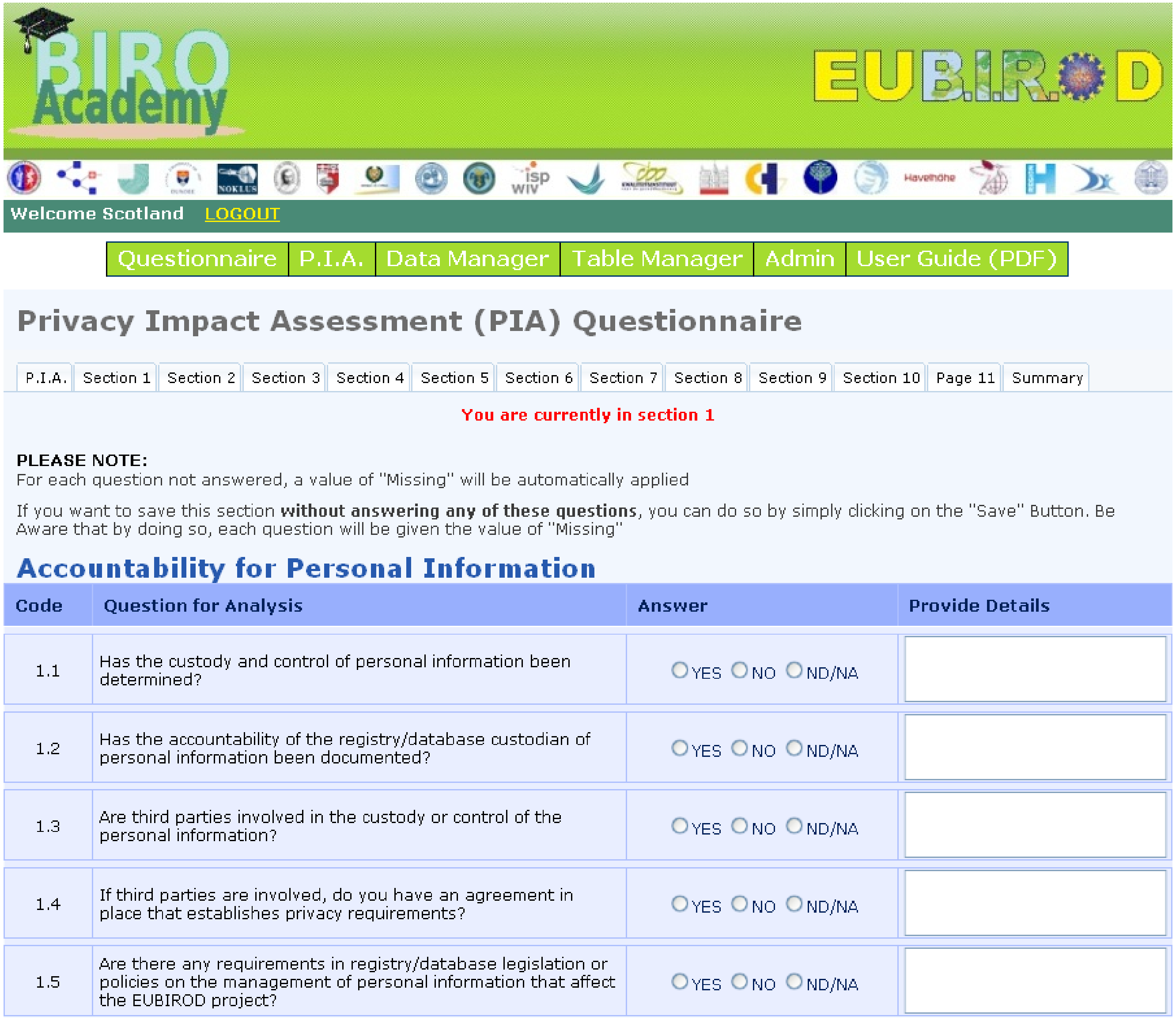
A8. Openness

A9. Individual Access to Personal Information

A10. Challenging Compliance

A11. Anonymization Process for Secondary Uses of Health Data

http://questionnaire.eubirod.eu



Privacy Impact Assessment (PIA) Questionnaire

You are currently in section 1

PLEASE NOTE:
For each question not answered, a value of "Missing" will be automatically applied

If you want to save this section **without answering any of these questions**, you can do so by simply clicking on the "Save" Button. Be Aware that by doing so, each question will be given the value of "Missing"

Accountability for Personal Information

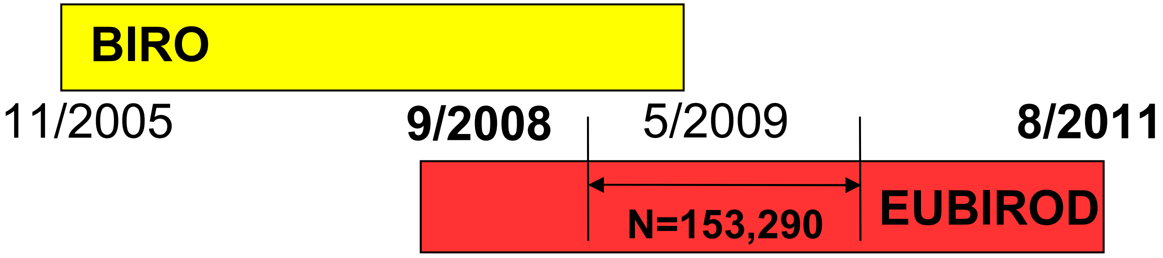
Code	Question for Analysis	Answer	Provide Details
1.1	Has the custody and control of personal information been determined?	○ YES ○ NO ○ ND/NA	
1.2	Has the accountability of the registry/database custodian of personal information been documented?	○ YES ○ NO ○ ND/NA	
1.3	Are third parties involved in the custody or control of the personal information?	○ YES ○ NO ○ ND/NA	
1.4	If third parties are involved, do you have an agreement in place that establishes privacy requirements?	○ YES ○ NO ○ ND/NA	
1.5	Are there any requirements in registry/database legislation or policies on the management of personal information that affect the EUBIROD project?	○ YES ○ NO ○ ND/NA	

Factors and the Scoring System

- The **scoring system** measures the level of compliance of local data processing with privacy principles according to an ordinal scale
  - **increasing factor score = increasing level of compliance**
- Scores are computed as a sum of responses to questions in each section, recoded either as 1 for a privacy protective conduct, or 0 for the opposite condition
- **To compare results across factors**, original values are presented as a percentage of the maximum attainable value (rescaled factors)
- **To compare results across registers**, the average of rescaled factors is used as a composite indicator of "**overall privacy performance**"
- Ad hoc R software has been developed for statistical analysis

EUBIROD Privacy Survey Sample (N=18)

University of Perugia (I)
Serectrix snc (I)
University of Dundee (GB)
Joanneum Research (A)
NOKLUS (N)
Paulescu Institute (RO)
University of Malta (M)
Republic of Cyprus (CY)
Sahlgrenska Institute (S)
University of Debrecen (H)
Institute of Public Health (B)
IDF (B)
Adelaide Meath Hospital (IRL)
CBO (NL)
Centre Hospitalier (LUX)
University of Ljubljana (SLO)
IMABIS Foundation (E)
Medical University Silesia (PL)
Havelhoe Hospital (D)
Hillerod University Hospital (DK)
Vuk Vrhovak University (HR)


BIRO

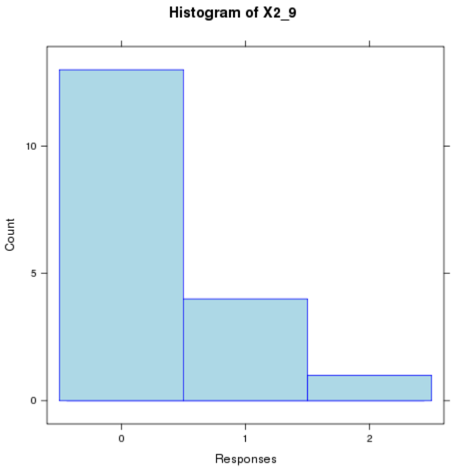
11/2005 **9/2008** 5/2009 **8/2011**

N=153,290 **EUBIROD**

Main Findings from Single Questions

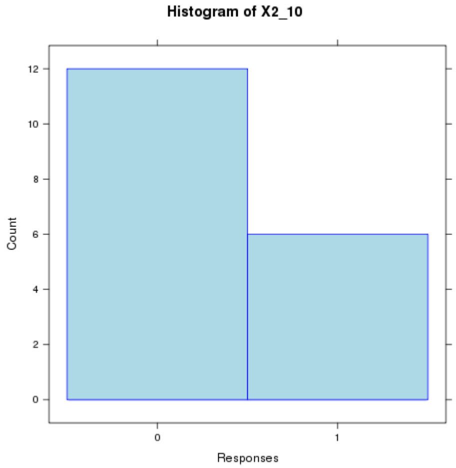
Responses to single questions highlight the following:

- diabetes registers normally don't have access to personal information from routine databases and/or multiple sources
- data linkage is performed only by half of the registries included in the survey
- the use of data for secondary purposes is hardly possible



Histogram of X2_9

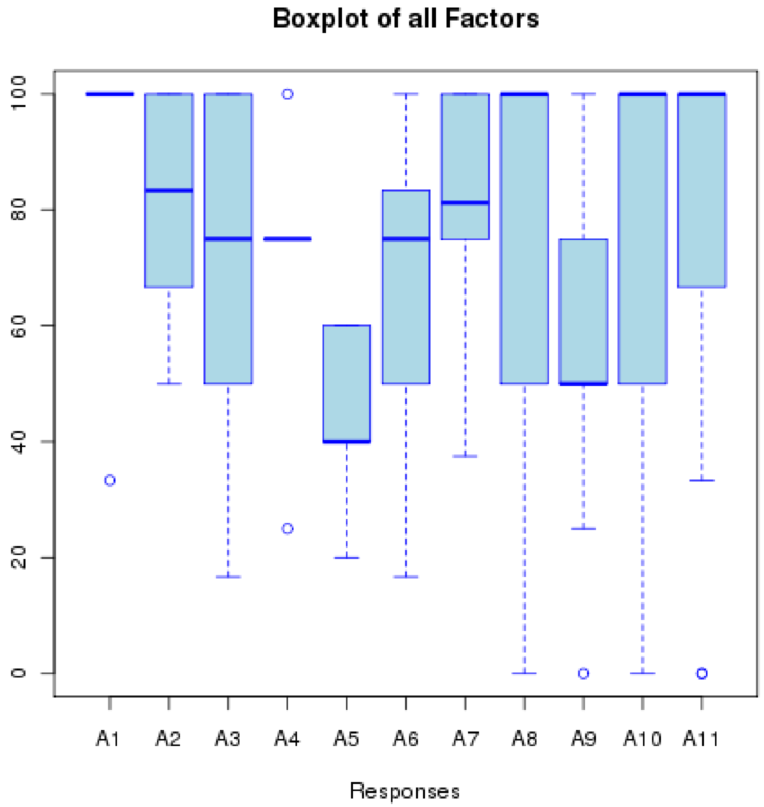
The possibility to collect some personal information from public databases is envisaged only in N=4 (22%) registries



Histogram of X2_10

Linking multiple sources through a common patient identifier is performed by N=6 (33%) registries

Standardized Comparisons of Factors Results



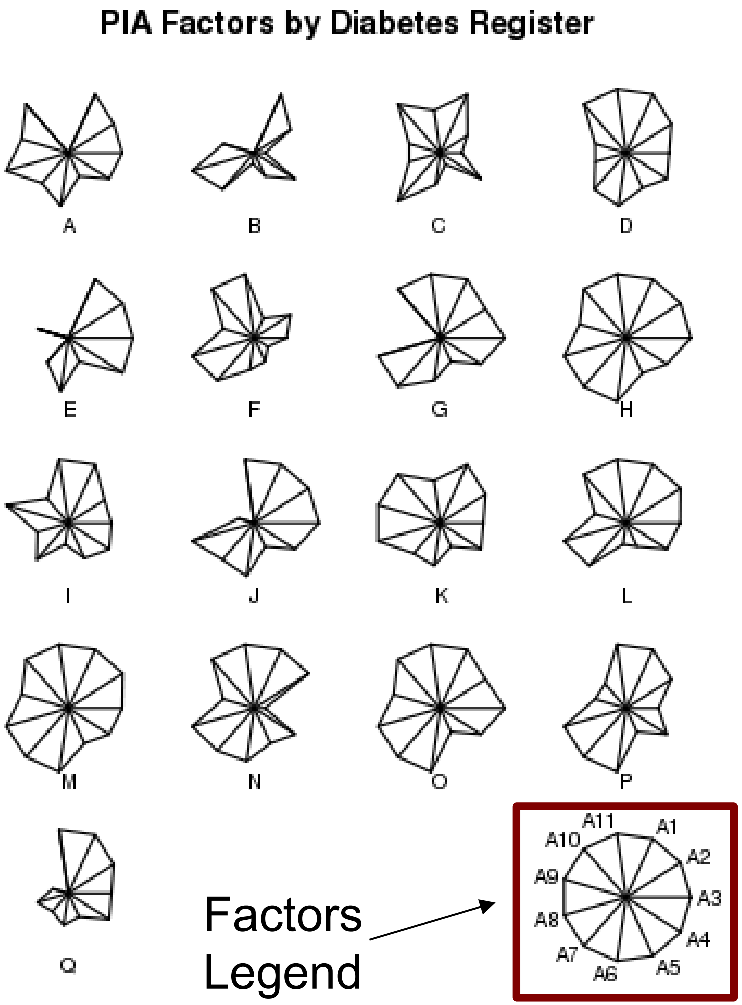
Low average (median):

A5: Disclosure and Disposition (40%)
A9: Individual Access (50%)
A3: Consent (75%)
A4: Use of Personal Information (75%)
A6: Accuracy (75%)

High Variability (standard deviation, range):

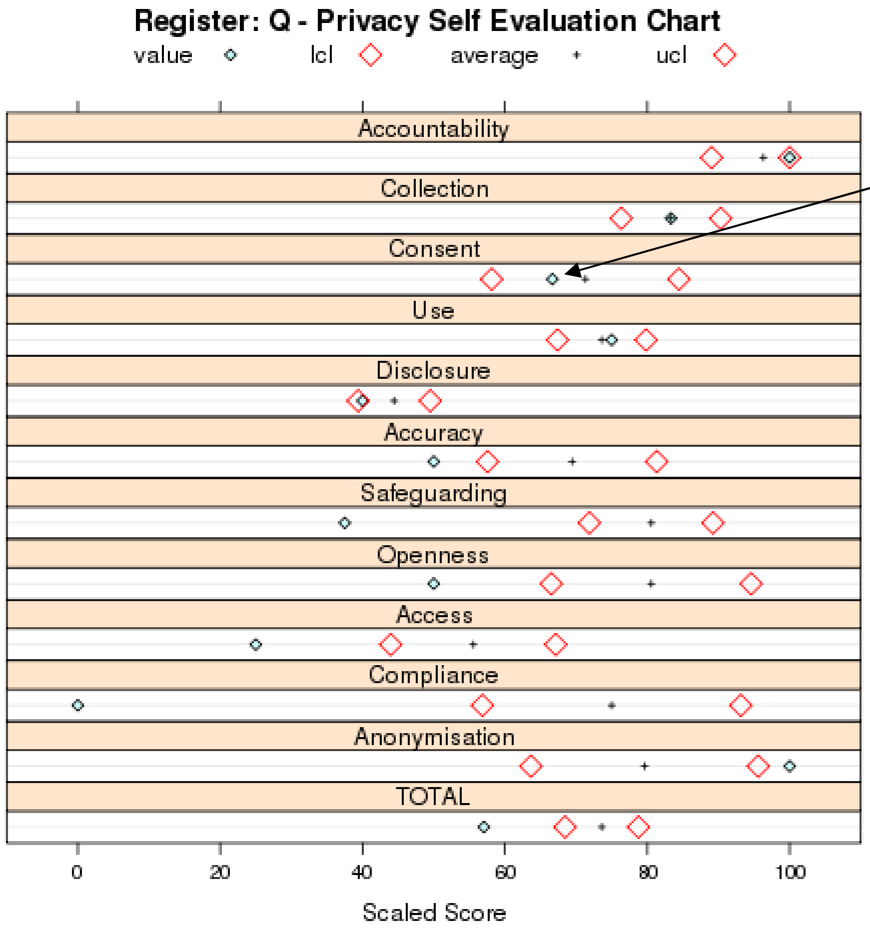
A10: Challenging Compliance (39%, 0-100%)
A11: Anonymisation (35%, 45-100%)
A8: Openness (30%, 0-100%)
A3: Consent (28%, 17-100%)
A6: Accuracy (26%, 17-100%)
A9: Individual Access (25%, 0-100%)

Analysis of Variability across Registers



- **<u>Starplots</u> summarize the "Privacy Profile" of each EUBIROD register included in the database**

Privacy Performance Self-Evaluation



- **For each factor and the overall score, each register can compare its <u>position</u>, against:**
  - **the 95% confidence interval around the average of the overall sample**
  - **the maximum attainable score (100%)**
- **The identity of centres is never disclosed**
- Example:
  - Maximum score in terms of accountability and anonymisation
  - Acceptable levels for collection, consent, use and disclosure
  - All other factors show poor privacy performance

Conclusions (1)

- **In several Member States, the balance between privacy protection and health research has been tipped in favor of the individual right to privacy. Only in few cases it is possible:**
  - to access personal information from routine databases and/or multiple sources
  - to perform data linkage
  - to use data for secondary purposes
- **Key areas of concern need targeted actions** to guarantee the right to privacy

Conclusions (2)

- The **Privacy Performance Self-Evaluation** methodology developed in EUBIROD can be used to **tailor specific corrective interventions at EU, National, Regional and Local level, based on explicit metrics**
  - the EU should provide Member States with legislation/guidelines that would ensure a sound interpretation of the Directive in public health applications
  - National, regional and local governments should foster the uptake of privacy principles/norms
  - The "**privacy performance self-evaluation tool" developed in EUBIROD** could be used to help managers of disease registers to enhance privacy protection and increase data accuracy and completeness

Final recommendation

- **A concerted action at both legislative and point of care levels is needed to achieve an optimal balance between the right to privacy and the right to the highest attainable level of health**