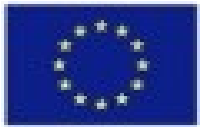


Directorate-General for Health & Consumers

Executive Agency for Health and Consumers

EUropean Best Information
through Regional Outcomes in Diabetes

WP5: DATA COLLECTION

D5.2: Privacy Impact Assessment

"*Privacy Impact Assessment Report*"

The EU.B.I.R.O.D. PIA Team

August 2010

This report is Deliverable D5.2 of "WP5: Data Collection, the European project "European Best Information through Regional Outcomes in Diabetes" (EUBIROD), co-funded by DG-SANCO, European Commission, 2008 (G.A. 2007115)

***Scientific Coordinator**: Prof.Massimo Massi Benedetti*
***Technical Coordinator**: Fabrizio Carinci*

A joint production of the EUBIROD Consortium:

Adelaide and Meath Hospital, Dublin, Ireland
Centre Hospitalier de Luxembourg, Luxemburg
Dutch Institute for Healthcare Improvement, Nertherlands
Havelhöhe, Berlin
Hillerød University Hospital, Hillerød, Denmark
IMABIS Foundation, Malaga, Spain
International Diabetes Federation, Belgium
Inst. Scient. Santé Pub. WIV, Brussels, Belgium
Joanneum Research, Austria
Medical University of Silesia, Katowice, Poland
Ministry of Health, Cyprus
NOKLUS, Norway
Paulescu Institute, Romania
Sahlgrenska Academy, Gothenburg, Sweden
Serectrix snc, Italy
University of Dundee, Scotland
University of Malta, Malta
University of Perugia, Italy
University of Debrecen, Debrecen, Hungary
University Children's Hospital, Ljubljana, Slovenia
Vuk Vrhovac University Clinic for Diabetes, Zagreb, Croatia

***WP Leader:** University of Dundee, UK*
***Task Leader:** Serectrix snc, Italy*

Compiled for Serectrix snc by:
Concetta Tania Di Iorio, Legal Expert, Serectrix snc, ITALY
Fabrizio Carinci, Senior Statistician, Serectrix snc, ITALY

The Privacy Impact Assessment Team:
Di Iorio Concetta Tania, Brillante Massimo, Adány Róza, Azzopardi Joseph, Battelino Tadej, Beck Peter, Boran Gerard, Cunningham Scott, de Beaufort Carine, Debacker Noemi, Jarosz-Chobot Przemyslawa Krystyna, Jecht Michael, Lindblad Ulf, Metelko Željko, Olympios George, Pruna Simion, Røder Michael, Skeie Svein, Storms Fred, Whiting David, Massi Benedetti Massimo and Carinci Fabrizio

Citation
Di Iorio CT, Adány R, Azzopardi J, Battelino T, Beck P, Boran G, Brillante M, Cunningham S, de Beaufort C, Debacker N, Jarosz-Chobot Przemyslawa K, Jecht M, Lindblad U, Metelko Ž, Olympios G, Pruna S, Røder M, Skeie S, Storms F, Whiting D, Massi Benedetti M and Carinci F, 0Privacy Impact Assessment Report, EUBIROD Consortium, 2010

Table of Contents

Executive Summary

The EUBIROD project aims to implement a sustainable European Diabetes Register through the coordination of existing national/regional frameworks and the systematic use of the BIRO technology. The project runs between 2008-2011, is co-funded by DG-SANCO, European Commission, and coordinated by the University of Perugia, Italy.

The project is based on the usage of the BIRO System, a tool specifically built to share an "Evidence-Based Diabetes Information System" among seven European countries. The system, developed between 2005-2009, has a structured architecture that involves two data processing steps, corresponding to a local and a global component, linked by a uni-directional flow of information.

The EUBIROD Privacy Impact Assessment (PIA) aims at documenting the impact on privacy of the BIRO System in the broader and more heterogeneous context of the EUBIROD Consortium, which includes nineteen partners maintaining large diabetes registers from different European countries.

Rolling out the system at this level involves testing its routine application in a framework where different management approaches may affect the completeness of the information contained in the registers and, consequentely, the comparability of results.

The "Privacy Impact Assessment" was planned as a specific task of work-package 5 (data collection) of the EUBIROD project, focusing on the following activities:

- identification of key elements of data protection in the management of diabetes registers
- classification of all components into factors/sub-factors relevant for privacy assessment
- creation of a questionnaire to collect information on data processing procedures
- analysis of the variability of approaches at the European level
- development of a platform to improve the management of privacy issues in disease registers

The fulfillment of these activities has allowed answering the following fundamental questions:

- how heterogeneous is the implementation of privacy requirements/principles among participating centres?
- which are the key areas of concern requiring advise and guidance?

The process involved the constitution of a dedicated team led by a legal expert.

Key elements of data protection (factors) have been identified by selecting, through a literature review performed in BIRO, privacy principles and norms that could be involved in data processing operations occurring in the management of diabetes registers.

A PIA questionnaire has been adopted to investigate the following:

- overall level of privacy protection
- heterogeneity in the implementation of privacy principles and requirements
- key areas of concern

The EUBIROD privacy impact assessment questionnaire is composed of a total of N=11 sections, one for each factor identified, including specific questions aimed at comparing data procedures in each register against privacy principles enshrined in EU and International legislation.

The following sections (factors) are included in the questionnaire: accountability of personal information, collection of personal information, consent, use of personal information, disclosure and disposition of personal information, accuracy of personal information, safeguarding personal

information, openness, individual access to personal information, challenging compliance and anonymisation process for secondary uses of health data.

Each section (factor) includes various questions (sub-factors) that allow drilling down into specific aspects of each factor included. For instance, section 2 ("Collection of personal information") includes: "Is personal information beeing collected directly from the individual? Have the purposes for which information is collected been documented? Are secondary uses contemplated for the information collected?

Possible answers include yes, no, not applicable or not determined, and an additional open field to provide comments, favouring a proper interpretation of responses and determining if assistance from the legal expert was needed.

The questionnaire was distributed to all partners at the first BIRO technical meeting (Rome, November 2009) as an empty Word document, to be filled in remotely with the collaboration of local database administrators, register managers and the task leader. Completed data were requested for the susequent meeting, scheduled after six months. During this timeframe, continuous legal advise has been provided to ensure the correctness and completeness of answers and to avoid any potential misunderstanding in the interpretation of the questions. Various rounds of submissions/corrections were needed to complete the process. A total of N=18 out of N=19 EUBIROD Centres were able to finalize the questionnaire (Special BIRO Academy Meeting, Rome, June 2010) and send it electronically to the Coordinating Centre of Perugia, Italy.

Each question has been coded as 0/1 according to an increasing level of compliance to privacy principles. Data were entered in an excel sheet and saved as a comma delimited text file.

A scoring scale has been created for each factor, summing up all relevant sub-factors and assigning equal weights to each component question. The possible range of the scores varies due to the different number of component questions. To compare the level of compliance across factors and registers, a standardized measure was derived by rescaling the values obtained as percentages of the maximum attainable score for each factor. The composite indicator of the overall privacy score attained by each register was computed as the average of the rescaled factors.

Statistical analysis included descriptive frequencies of factors and overall scores, medians, means and their associated 95% confidence intervals, plus a range of exploratory graphical outputs. An "ad hoc" software written in the R statistical language has been specifically developed for the scope.

Results of the EUBIROD privacy impact assessment successfully allowed to investigate the degree of heterogeneity among participating registers, identifying key areas of concerns in the implementation of privacy principles across Europe.

Responses to single questions highlight the following:

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

The analysis of factor scores shows that the major areas of concern (median, range) are: disclosure and disposition of personal information (40%, 20-60%) and individual access to personal information (50%; 0-100%). The following factors are also highly problematic: consent (75%; 17-100%), use of personal information (75%; 25-100%), accuracy of personal information (75%; 17-100%).

Factors showing a high variability (standard deviation) include the following: challenging compliance (39%), anonymisation (35%), openness (30%), consent (28%), accuracy of personal information (26%), individual access to personal information (25%).

The range of overall scores achieved by EUBIROD registers was 69-79% (mean:74%, standard deviation: 11%), with a median close to 75% and almost 20% of the sample falling above 80% of the maximum performance.

The continuous application of the EUBIROD privacy impact assessment methodology would allow monitoring these results in direct interaction with end users. To this end, an online platform has been specifically developed to conduct a continuous “privacy performance self-evaluation”.

The web platform includes an electronic version of the questionnaire and a management system that allows its applicability to new users. New questionnaires submitted to the system shall be validated by the task leader. Validated files are automatically translated into coded data and directly submitted to the R routine that updates results on the server. Graphical outputs summarizing the scores achieved by each centre are fed back to the relevant user to allow comparing the profile of own practice against that of the entire sample. All outputs are provided in anonymous format, i.e. no centre is able to identify results relative to other centres.

The EUBIROD privacy survey has delivered accurate information on the level of privacy protection in eighteen diabetes registries across Europe, allowing to analyse how the Data Protection Directive has been implemented in this particular field.

The findings of our survey could be used to develop targeted actions at both European and National levels. To ensure that accurate public health monitoring is not jeoupardized by legislative constraints, the EU should provide suitable guidelines to Member States specifying how to deal with crucial issues e.g. data accuracy, consent, data linkage, secondary uses of health information, etc. At the same time, National governments should foster the uptake of privacy principles locally. The “privacy performance self-evaluation” tool that has been conceived in EUBIROD may represent a suitable means to achieve such an objective at national, regional and single centre level.

The attainment of an optimal balance between public health needs and privacy requires that ethical values enshrined in the EU and international legislation are fully understood and duly applied across Europe. If realized, it would enhance the accuracy and completeness of data collected across Europe.

The adoption of appropriate safeguards in data processing operations that pose privacy risks should be fostered. For instance, computerized data linkage, a practice that is increasingly posing privacy concerns, can be safely performed using best approches, e.g. trusted third parties, which can guarantee the respect of privacy norms.

The “privacy performance self evaluation platform” conceived in EUBIROD may represent an example of how privacy enhancing registers can be collaboratively fostered. Quality improvement schemes, rather than “privacy league tables”, are definitely needed to increase both the level of privacy protection and the information content of available clinical databases.

Concerted actions at both the legislative and implementation level should be promoted to achieve the right balance between the right to privacy and the right to the highest attainable level of health.

1. Introduction

The influential role of information technology in all aspects of modern society provides enormous opportunities for the implementation of innovative applications in the health sector. The availability of health data anytime, anywhere, allows for health systems to be optimised and to respond to the particular needs of citizens most effectively and efficiently. Through systems integration, medical records relative to a single individual may be stored at different sites and conveyed to a practitioner to support an informed judgement on the best therapeutic option according to the most relevant guidelines. At the same time, policy makers at all levels may better understand how to improve services through routine monitoring of their performance and the availability of up-to-date indicators computed through computerized data linkage.

Despite of an overwhelming amount of individual data recorded on a routine basis, health information across Europe is still fragmented, under-utilized, insufficiently summarized for the needs of policy makers[1].

On a global scale, there is a need to improve access to high quality information for health policy research. To this end, and to make health systems more sustainable in a situation of financial pressure, many different sources of information need to be linked and integrated. However, sharing highly sensitive data poses fundamental ethical questions that cannot be underestimated and should be faced by citizens, health professionals, health care organizations, and policy makers.

The role of governmental institutions in protecting public health should be to create the conditions for the above stakeholders to overcome the existing barriers in the public interest. The European Commission, with its ambitious goals in the fields of research and realization of the health strategy, should support Member States in fulfilling this task through targeted legislative instruments.

The privacy and ethical values of citizens must be safeguarded to avoid improper usage of personal data. However, the respect of privacy should not be invoked above its proper interpretation and unnecessarily limit the free flow of information across countries, a principle enshrined “per se” in both the EU Data Protection Directive[2] and the EU Treaties[3].

1.1 Data Protection Legislation in the EU

Of all the human rights in the international catalogue, the **right to privacy** is perhaps the most difficult to define[4].

Definitions of privacy vary widely according to contexts and environments. Nevertheless, privacy is usually seen as the way of drawing the line of how far a society can intrude into a person’s private life.

Privacy has been defined as the “right to be left alone”[5]; or “the right of the individual to be protected against intrusion into his personal life or affairs, or those of his family, by direct physical means or by publication of information[6].

Although there is no unique definition of privacy, it has been recognized as a fundamental human right in many international instruments.

The 1948 *Universal Declaration of Human Rights* has been the first international binding instrument to recognise privacy as a human right, specifically protecting territorial and communication’s privacy[7]. Art. 12 states: *“No one should be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks on his honour or reputation. Everyone has the right to the protection of the law against such interferences or attacks”.*

In addition, numerous international human rights treaties specifically recognize privacy as a right. The International Covenant on Civil and Political Rights (ICCPR – Art. 17)[8]; the UN Convention on Migrant Workers (Article 14)[9], and the UN Convention on Protection of the Child (Art. 16)[10] adopt the same language. On the regional level, various treaties make these rights legally enforceable.

For instance, Art. 8 of the *European Convention for the Protection of Human Rights and Fundamental Freedoms* (1950)[11] states that *"Everyone has the right to respect for his private and family life, his home and his correspondence. There shall be no interference by a public authority with the exercise of this right except as in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health of morals, or for the protection of the rights and freedoms of others".*

The Convention has created the European Commission of Human Rights and the European Court of Human Rights to oversee enforcement. Both have been active in the enforcement of privacy rights, and have consistently viewed Article 8's protections expansively and interpreted the restrictions narrowly[12].

The Court has reviewed Member States' laws and imposed sanctions on numerous countries[13]. It also reviewed cases of individuals' access to their personal information in government files to ensure that adequate procedures exist[14]. The interest towards the right of privacy increased in the 1960s and 1970s with the advent of information technology.

The availability of powerful computer systems has increased the demand for specific rules governing the collection and handling of personal information.

Two crucial international instruments in the evolution of data protection are the Council of Europe's "Convention for the Protection of Individuals with regard to the Automatic Processing of Personal Data"[15] and the Organization for Economic Cooperation and Development's (OECD) "Guidelines Governing the Protection of Privacy and Trans-border Data Flows of Personal Data"[16].

The above agreements set out specific rules covering the handling of electronic data, describing personal information as data that have to be protected at every step: from collection to storage and dissemination. They exerted a profound effect on the enactment of laws around the world. Nearly thirty countries have signed the COE Convention. The OECD guidelines have been widely used in national legislations, even outside the OECD member countries. The development of privacy protection in the EU took a step forward with the Council of Europe Convention on Human rights and Biomedicine (Oviedo 1997), which reinforced the principles that everyone is entitled to the right to privacy and confidentiality of personal medical data and the right to be informed about his/her health[17].

Finally, the *Charter of Fundamental rights of the European Union* (2000/C 364/01)[18] specifically provided protection of personal data. Art. 8 states: "*Everyone has the right to the protection of personal data concerning him or her. Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified. Compliance with these rules shall be subject to control by an independent authority".*

The Charter of Fundamental Rights was fully incorporated in the *European Constitution* (forming its part II)[19], signed in Rome on the 29th of October 2004. Although the Parliament, the Council and the Commission solemnly proclaimed the Charter on the 8th of December 2000, the Charter is not part of the Union's Treaties and therefore it has no binding legal force. The Constitution achieved a major breakthrough, which allows the Union to have its own catalogue of rights, binding for all European countries and enforceable through the Court of Justice, which will in fact ensure that the Charter will be adhered to. At present, the Treaty of Lisbon[20] also guarantees the enforcement of the Charter of Fundamental Rights. The EU therefore acquires a catalogue of civil, political,

economic and social rights, which will be legally binding also for Member States as regards the implementation of Union law.

It is worth noting that the content of the Charter is broader than that of the European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR)**,** signed in Rome on 4 November 1950 and ratified by all the Member States of the Union.

Whereas the ECHR is limited to civil and political rights, the Charter of Fundamental Rights covers other areas such as the right to good administration, the social rights of workers, the protection of personal data and bioethics.

Finally, the *Additional Protocol to the Convention on Human Rights and Biomedicine, concerning Biomedical Research* (2005)[21] further reinforced the duty of confidentiality in the handling of personal information in health research and reaffirmed the obligation to treat them according to the rules relating to the protection of private life.

In line with all the aforementioned instruments, the EU has adopted a privacy legislative model embracing comprehensive laws. The model is based on a general and abstract law that governs all aspects of the handling of personal information: from collection to use and dissemination, by both the public and private sectors.

The 1995 Data Protection Directive (95/46/EC)[2] sets up a common level of privacy among European countries, ensuring compliance through the establishment of a regulatory body.

The Directive reinforced current data protection laws and established a range of new rights and basic principles, namely: the right to know where the data originated, the right to have inaccurate data rectified, a right of recourse in the event of unlawful processing, and the right to withhold permission to use data in some circumstances. The Directive contains strengthened protections over the use of sensitive data.

Art. 7 of the Directive establishes a set of criteria of *"legitimate processing".* Processing, in order to be legitimate, has to take place: either with the unambiguous consent of the data subject, or where this is necessary for the performance of a contract with the data subject, for compliance with a legal obligation, or for the performance of a government task, etc.

More stringent conditions apply to the processing of *sensitive data*, such as medical data. Here, the processing of sensitive data is considered, in principle, not legitimate and Member States have to prohibit their processing, unless special conditions verify.

According to Art. 8, the processing of sensitive data is allowed when:

- the data subject has given his explicit consent to the processing of those data, or
- processing is necessary for the purposes of carrying out the obligations and specific rights of the controller in the field of employment law in so far as it is authorized by national law providing for adequate safeguards; or
- processing is necessary to protect the vital interests of the data subject or of another person where the data subject is physically or legally incapable of giving his consent; or
- processing is carried out in the course of its legitimate activities with appropriate guarantees by a foundation, association or any other non-profit-seeking body with a political, philosophical, religious or trade-union aim and on condition that the processing relates solely to the members of the body or to persons who have regular contact with it in connection with its purposes and that the data are not disclosed to a third party without the consent of the data subjects; or
- the processing relates to data which are manifestly made public by the data subject or is necessary for the establishment, exercise or defence of legal claims.

Importantly, the general prohibition of Art. 8(1) shall, according to Art. 8(3), also not apply where the data are required: for the purposes of preventive medicine, medical diagnosis, the provision of care or treatment or the management of health-care services, and where those data are processed by a health professional subject under national law or rules established by national competent bodies to the obligation of professional secrecy or by another person also subject to an equivalent obligation of secrecy.

Moreover, Member States may, according to Art. 8(4), for reasons of substantial public interest, lay down exemptions, in addition to those laid down, either by national law or by decision of the supervisory authority.

Art. 8(3) is extremely important for the health sector, since it justifies the collection, use, and processing of health data, for the specified purposes, without the patient's consent. Although the free and informed consent will be necessary if, for instance, those data would be further used for research purposes or any other secondary use.

Recitals 33-34 of the Directive are of utmost importance for the rightful interpretation of Art. 8(3-4) of the Directive. Recital 33 explains that "*derogations must be explicitly provided for in respect of specific needs, in particular where the processing of these data is carried out for certain health-related purposes by persons subject to a legal obligation of professional secrecy*". Furthermore, Recital 34 explicitly identifies public health, social protection, scientific research and government statistics as reasons of public interest that justify derogation to the prohibition of processing sensitive data, save that specific and suitable safeguards are provided to protect the fundamental rights and the privacy of individuals.

Other derogations, relative to information to be given to data subjects and access, are also envisaged for statistical and scientific research in Art. 11-12 of the Directive. The reference to professional secrecy contained in Art. 8(3) and Recital 33, is crucial for obtaining a more effective protection of privacy in the handling of sensitive health data. Although issues surrounding the confidentiality of health data are not fully covered by the Directive, the reference made to the obligation of confidentiality in the Directive represents a step forward towards the possible harmonization of European legislations. At least, it imposes to Member States, in a binding form, the *duty of confidentiality* to any person involved in the processing of personal sensitive data, such as health data.

The duty of confidentiality originates from the duty of professional secrecy incumbent on health professionals either through a law or code of conduct. The principle of confidentiality of medical information, derived by the Hippocratic Oath, can be considered one of the oldest principles applying to data protection. Although privacy and confidentiality are conceptually distinct, they are strictly interrelated and need to be consistently implemented among European countries in order to enhance the protection of privacy when sensitive data are involved: in this regard, confidentiality could be conceived as a means to protect the right to privacy.

In order to conduct scientific research without falling under the binding rules of the Directive, data should be rendered anonymous. *Recital 26* states that: *"principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable".*

Art. 2 of the Directive specifies that an *"identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity"*. In order to determine whether a person is identifiable, "*account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said person".*

When the data subject could be identified with reasonable means (directly from the data itself or indirectly through the combination of other means), data cannot be considered anonymous and, therefore, fall under the Directive principles[22], including the need to obtain expressed consent from the data subject.

However, the identification of the data subject through *"reasonable means"* is a vague concept. In each particular case, reference to the state of the art in decoding and/or other similar techniques should be made to indirectly assess what "reasonable means" stands for. The definition involves an "ad hoc" evaluation of the likelihood of re-identification based upon technical matters[23]. Data could be then considered anonymous when *"it would be reasonably impossible for the researcher or any other person to re-identify the data"*[24].

In such a case, the interest of the data subject to maintain his/her data private and confidential is protected "ipso iure" by anonymisation, rendering the processing legitimate even without consent.

Accordingly, data processed anonymously for research purposes should be regarded as falling outside of scope of the Directive whenever no direct/indirect identification is possible with reasonable means, according to the state of the art.

Recital 26 of Directive also fosters the development of codes of conduct, within the meaning of article 27, to provide guidance on the ways in which data may be rendered anonymous; thus, leaving the definition of such a crucial issue at the implementation of the Directive at national level.

Importantly, the Directive imposes an obligation on Member States to ensure that personal information related to EU citizens has the same level of protection when is exported to, and processed in, countries outside the EU. As a result, countries refusing to adopt adequate privacy protections may be unable to conduct certain types of information flows with Europe, especially when the transmission of sensitive data is involved.

In line with the Directive, in 1997 the Council of Europe enacted a Recommendation on the Protection of Medical Data[25] acknowledging that medical data require even more protection than other non-sensitive data, and reaffirming that the respect of rights and fundamental freedoms, in particular the right to privacy, has to be guaranteed in the collection and processing of medical data.

Therein the processing of medical data is in principle prohibited, unless appropriate safeguards are provided by domestic law.

One of such safeguards is that only health-care professionals, bound by rules of confidentiality, should process medical data; though persons acting on their behalf are also allowed to perform the same duties if subject to the same or similar rules.

According to the Recommendation, medical data may be collected, from the data subject or from other sources, if permitted by law, for public health reasons [Principle 4.3(a)] and for the purposes listed in Principle 4.3(b):

- for preventive medical purposes or for diagnostic or therapeutic purposes (in this case data may also be processed for the management of medical service operating in the interest of the patient);
- to safeguard the vital interests of a data subject;
- to respect specific contractual obligations;
- to establish, exercise or defence in a legal claim.

Thus, the Recommendation reaffirms and strengthens the rules set forth by the Directive.

Medical data may be collected without consent "for the purposes of" (i.e., in the interest of) public health, including the management of health services. For health research, the processing of health data is considered legitimate whenever data are rendered anonymous, with techniques being continuously updated and kept efficient.

Accordingly, health data handled for research purposes must not be published in a form that enables data subjects to be identified, unless data subjects have given their consent for publication and/or specifically permitted by domestic law.

1.2 The EUBIROD Project

"EUropean Best Information through Regional Outcomes in Diabetes" (EUBIROD)[26] is a three year public health project in the field of diabetes started on the 1st September 2008, sponsored by the European Union under the Health Information Strand of the Public Health Program (DG-SANCO).

The project mission is to implement a sustainable European Diabetes Register through the coordination of existing national/regional frameworks and the systematic use of the BIRO technology.

The system fosters the objectives of the Conclusions of the EU Council for the systematic data collection and monitoring of diabetes complications and health outcomes across Europe. EUBIROD proposes an action to implement, extend, and customise the application of the BIRO technology in 20 States, including EU Member States, Acceding/Candidate Countries, and EFTA Countries. Participants will be connected through a system that will safely collect aggregated data and produce systematic EU reports of diabetes indicators, which will be used to develop recommendations for policy makers.

The project includes nineteen partners managing diabetes registers in different European regions, one technological partner leading the privacy impact assesment presented in this report, one collaborating institution from outside Europe, and a major representative of the needs and expectations of people with diabetes, the International Diabetes Federation.

The main expected output of the project is the production of the "European Diabetes Report": an analysis of quality of care and outcomes in diabetes based on standardized criteria over a reference population of 500,000 subjects.

The project supports improved information at both the micro and the macro levels: it facilitates activities for planning and management of diabetes care in regional health systems, delivering information that is directly applicable at the Community level by European institutions.

1.3 The BIRO System

The BIRO project[27] built a "Shared Evidence-Based Diabetes Information System" (SEDIS) among seven European countries. The system has a structured architecture that involves two data processing steps, corresponding to a local and a global component, linked by a uni-directional flow of information (Figure 1.1).

A basic version of the system runs in each single register ("local SEDIS") to produce initial estimates for the local population. All partners in the network, using the same standardized procedures, repeat the process at their best convenience. Regional estimates are then sent to a central server that compiles "partial" results into a European report ("global SEDIS"). A web portal delivers user-friendly information for local registers.

Functionality of the system is ensured by three fundamental elements: a concept and data dictionary including standardized evidence-based definitions in XML format; a report template to structure presentation of end results; and statistical methods required to produce them.

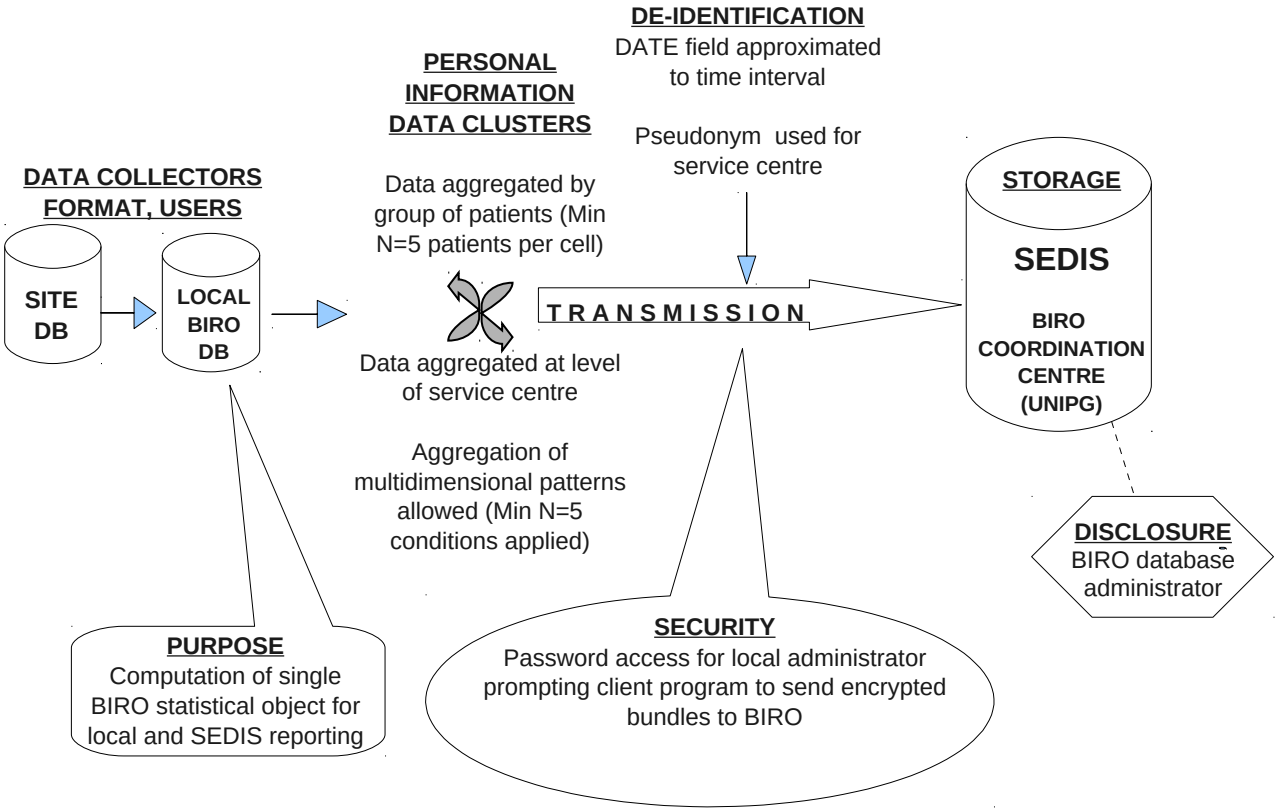
The same structure is used to automate the production of BIRO reports for individual centres and the whole network.

The data model includes a BIRO XML export, loaded by a Java-powered "Database Manager" into a local (Postgres) database that is directly accessed by R statistical routines to produce aggregate results. "Statistical objects" are defined as "elements of a distributed information system carrying essential data in the form of embedded, partially aggregated components, that can be used to compute a summary measure or relevant parameter for the whole population from multiple sites".

Communication software is used to send statistical objects to a central server, where an ad hoc Java Importer loads them into a central BIRO database, and a global repository is maintained.

Functions are used to process aggregate data submitted by local registers until a global pooled estimate is produced and published in pdf and html format on a dedicated web portal.

Figure 1.1: BIRO System Architecture



1.4 Privacy Analysis of the BIRO System

Privacy impact assessment of the BIRO system has been described in detail elsewhere[28]. Here follows a brief description of its major findings.

The BIRO system involves medical records collected by diabetes registries at national or regional level, processed to support benchmarking and public health monitoring at the international level.

In this case, local data processing is subject to Art. 8(3) of the EU Directive[2]: each centre collects information related to an identified or identifiable natural person for the purpose of setting up diabetes registries. Hence, data could be considered collected and processed for purposes of preventive medicine, medical diagnosis, the provision of care or treatment or the management of healthcare services. In this case, the data collector is exempted from requesting consent from the data subject, in consideration of the need to protect the competing and general interests of societies in improved healthcare. The further processing of these data, other than caring for the patient and managing health services, would not be covered by the exemptions of Art. 8(3): in other words, consent would be required for any secondary use of those data.

However, according to Art. 11(2), for research and statistical analysis, even if consent was required in the first instance, the provision of information to the data subject could be waived if it proves impossible or would involve a disproportionate effort.

The exemptions provided by the Directive are in line with the principles contained in the Convention on the Protection of Individuals for the Automatic Processing of Personal Data (1981)[15], envisaging the possibility of restricting the exercise of the data subject's rights with regard to data processing operations that pose no risk [Art. 9(3)]. Examples of no-risk or minimal-risk operations are therein considered, in particular, the use of data for statistical work, if those data are presented in aggregate form and stripped of their identifiers, as in BIRO. Similarly, scientific research is included in this category.

In terms of data transmission, BIRO centres send only aggregate records to the central server. For the most sensitive variables, aggregated records are not transmitted if groups contain less than five patients. Statistical objects are sent as tables stored in compressed bundles of flat text comma delimited files (CSV). Hence, there is no possibility, either directly or indirectly, that a patient could be identified with "reasonable means".

In broad terms, the disclosure of information related to clinical centres or individual professionals may also pose particular privacy concerns. The Consortium felt that this factor could jeopardize the level of data sharing and eventually discourage participation to the project.

The issue raises an interesting point that may constitute a future area of contention: disclosing information on small centres may lead, without unreasonable means, to the identification of doctors and, eventually, of individual patients. In addition, it could imply judgements on individual centres' performance.

In consideration of the above concerns, Centres' IDs have been protected through the use of a pseudonym, together with a reporting system based on percentages rather than absolute numbers. Accordingly, the size of single Centres would be hidden, avoiding their indirect identification by third parties.

Aggregated statistical objects are sent to the central statistical engine to carry out global analysis.

A communication software has been specifically developed to ensure secure information exchange between the regional systems and the central SEDIS. To facilitate secure data transmission in BIRO, modern technologies have been selected and successfully used, complying with security requirements enshrined in both the EU and international data protection norms.

Global reporting does not pose any direct or indirect risk to privacy, as anonymous data sent by BIRO centres is transmitted to SEDIS in a secure environment and further processed in aggregate form.

The last issue relates to trans-border data flow: the central database is located outside national boundaries. The BIRO System, as already demonstrated, processes only anonymous data; therefore, privacy rules should not limit its implementation.

Nevertheless, the free flow of information, regardless of frontiers, is also a principle enshrined in Art.10 of the European Human Rights Convention[29]. Accordingly, Art.12 of the Convention on the Protection of Individuals with regard to Automatic Processing of Personal Data (1981) and Art.25 of the Directive discipline the transborder data flow.

The main rule contained in Art.12(2) of the Convention, is that, in principle, obstacles to trans-border data flows are not permitted between Contracting States in the form of prohibitions or special authorisations of data transfers. The rationale for this provision is that all Contracting States, having subscribed to a common core of data protection provisions set out in Chapter II, offer a certain minimum level of privacy protection.

In addition, no restrictions should be placed on the trans-border flow of medical data towards a State that has not ratified the Convention when the protection of medical data can be considered to be in line with the principle of equivalent protection therein laid down.

Therefore, the EU Directive allows the cross border flow of personal data only when an adequate level of privacy protection is envisaged in the countries involved in the processing operations.

Consistently with the interpretation of the Convention, countries that have fully implemented the Directive are automatically allowed to execute trans-border data flows: complying with the Directive ensures, “ipso iure”, an adequate level of protection.

BIRO centres belong to European countries that have fully implemented the EU Data Protection Directive and ratified the Convention. Hence, an adequate level of privacy protection is fully guaranteed across those countries. It follows that the exchange of data envisaged in the BIRO project is legally viable, considering the system architecture and the composition of the Consortium.

In accordance with EU and International legislation, reports will never allow either the data subjects or the local centres to be identified.

Potential privacy risks in the usage of the BIRO system are summarized in Table 1.1, showing the nature and level of identified risks, along with the required mitigation strategies.

Technological solutions have been duly implemented in B.I.R.O. taking into account such potential weaknesses.

In conclusion, as far as the EU legislation is concerned, data processing occurring in BIRO is to be considered legitimate, although domestic laws may provide more stringent rules to be specifically examined in each case.

Table 1.1 Privacy Contingency Risks

Element	Nature of risks	Level of risks			Comments	Mitigating Mechanisms
		Low	Medium	High		
Individual data: Pseudonym used for patients' IDs + Data is Aggregated (N=5 patient per cell)	Individual privacy	X			Pose an indirect risk to individual's privacy	Non-Reversible De-identification
Pseudonym used for Centres IDs	Non-Individual Privacy	X			Pose an indirect risk to Centres' privacy	Non-reversible De-identification
Data Transmission	Security Measures	X			Pose an indirect risk to individual's privacy	Encryption
Access to the BIRO network	Security Measures		X		Pose an indirect risk to individual's privacy	Secure applications Hacking tests
Global Statistical Analysis	Individual privacy + Non-Individual Privacy + Security Measures	X			Pose an indirect risk to individual's privacy and centres privacy	Non-reversible de-identification + Encryption

1.5 EUBIROD Privacy Impact Assessment

The EUBIROD Privacy Impact Assessment (PIA) aims at documenting the impact on privacy of the BIRO Health Information System in the wider and more heterogeneous context of the EUBIROD consortium, including diabetes Centres from nineteen European countries.

Rolling out the system at a European level involves encompassing different approaches that may impact on data completeness and comparability of results.

The "Privacy Impact Assessment" in the EUBIROD project focuses on:

- Identification of the key elements of data protection in the management of diabetes registers
- Definition of main factors in the evaluation of privacy issues
- Creation of a targeted tool (questionnaire) to collect data on procedures used across the EUBIROD Consortium
- Analysis of the variability of approaches at the European level
- Adoption of a targeted tool to improve the management of privacy issues: Privacy Performance Self-Evaluation of disease registries

The fulfillment of the above activities will allow answering crucial questions such as:

- how heterogeneous is the implementation of privacy requirements/principles among participating centres?
- which are the key areas of concern requiring advice and guidance?

2. Materials and Methods

2.1 Target population of diabetes registers

The target population of the EUBIROD privacy impact assessment includes a sample of N=19 registers that have been extensively documented in the scientific literature. Here follows a description of the N=18 registers that actively collaborated and contributed with primary data to the preparation of the present report.

In **Austria**, the Healthgate information system developed by Joanneum Research started in 1998 under the aegis of the Styrian Provincial Government, providing funding for the development of an information system for quality management in diabetology. Today, it covers different regions of Austria and Germany, offering data management and online benchmarking for disease management programs under the banner of the Austrian and German FQSD initiative (Forum Qualitätssicherung in der Diabetologie) [30]. The majority of data are voluntarily made available by FQSD members. A total of 2,600 patients per year from the Styria region will be included in the EUBIROD reports.

In **Belgium**, the Scientific Institute for Public Health runs the Initiative for Quality Promotion and Epidemiology in Diabetes (IQED) since 2001, with the participation of all secondary care diabetes centres (n=120) covering a population of almost 100.000 diabetes patients on > 2 ins inj/day. The data collection is scheduled every 18 months, based on DiabCare information sheet, followed by feedback and report. Sentinel network practices have been also investigated to assess the quality of primary care in Belgium[31]. Annual EUBIROD reports will include nearly 11,000 patients from the national survey.

In **Croatia**, the National Diabetes Registry[32], was founded to enable promotion of diabetes care, assessment of the prevalence and incidence of diabetes and its acute and chronic complications, and follow-up of morbidity, mortality, and basic clinical indicators. All primary and secondary care physicians who provide diabetes care are obliged to supply data on their patients to the Vuk Vrhovac University Clinic for Diabetes, Endocrinology and Metabolic Diseases on an annual basis. By the end of 2007 there were more than 80,000 patients registered in CroDiab. Over 40,000 subjects registered in CroDiab will be included in EUBIROD annual reports.

In **Cyprus**, with the opening of the first multidisciplinary diabetes clinic on the island at Larnaca Hospital, a new system for data collection has been introduced to underpin the creation of a National Diabetes Register. The electronic register was created in April 2007. It was developed in Microsoft access by the Ministry of Health, based on the BIRO common dataset, with some additional data being collected and recorded. All the data collected until August 2008 have been entered with nearly 900 active patients currently included in the database. The expansion of the system to the entire island has been planned to occur during the next three years. Annual EUBIROD reports will include over 800 patients per year registered in the Larnaca clinic.

In **Denmark**, a total of 2,500 active patients followed up by the Hillerod Hospital will be included in EUBIROD reports.

In **Germany**, the diabetes register of Rheinland-Pfalz, a region with a total population of over 4,000,000 inhabitants, participates to the German FQSD initiative. The system collects data on N=100 participating centres, with over 14,000 active patients followed up every year. Currently, only diabetologists are connected to the system. However, the plan envisages its expansion to include linked data from primary care and diabetes management programs. Nearly 15,000 patients will be included in EUBIROD reports.

In **Hungary**, the General Practitioners' Morbidity Sentinel Stations Program (GPMSSP) was launched by the School of Public Health, University of Debrecen, Hungary and the National Public Health and Medical Officer Service in May 1998. Data collection started at the beginning of October 1998. Four counties involved in the program were chosen to represent the eastern and western parts of Hungary. Direct information can be obtained about the morbidity of the selected diseases in the study population. Since 2008 the reporting system is fully computerised via a web-based application. The monitoring system also serves as a research infrastructure for epidemiological investigation[33]. A total of 1,400 patients per year will be included in EUBIROD reports.

In **Ireland**, the diabetes centre at the Adelaide and Meath Hospital Dublin is comprised of outpatient, research and day care functions. The centre serves the Dublin region of Tallaght, Clondalkin, Firhouse, Rathfarnam, Terenure, Templeogue, along with West Wicklow and parts of Co.Kildare. To help manage patient care, the centre operates an electronic diabetes patient database with currently over 6,000 active patient records. The Diamond diabetes database has been in place for several years. A total of 4,000 subjects per year are planned to be included in EUBIROD reports.

In **Italy**, the Department of Medicine at the University of Perugia has been performing its research and clinical activity in the field of diabetes since the early 70s. The Umbria Diabetes Register started in 1999 with grants assigned by the Italian Ministry of Health, to cover all major diabetologic centres in the region through an innovative platform for data exchange[34]. In 2007, a data linkage project sponsored by the Regional Department of Health allowed to integrate this information with administrative databases, allowing to capture nearly 68,000 diabetic patients across the region. The Umbria database will contribute to EUBIROD with over 10,000 active patients per year.

In **Luxembourg**, the incidence of type1 diabetes mellitus in children has been studied since the end of the eighties. In order to obtain more information on the other core and secondary indicators, a collaborative action was undertaken, financially supported by the Ministry of Health, Luxembourg

and coordinated by the CRP santé. Data were collected using two different sources. The medical administrative source, the UCM (reimbursement structure) provided data for several outcome and process indicators. No clinical data were accessible through this source. A questionnaire based data collection was proposed to all doctors involved in diabetes treatment in Luxembourg. Questionnaires (based on DIABCARE), were filled out by the doctors, on all patients with diabetes, seen in their outpatient clinic over a 6 months period of time. Data available through this method contributed to the EUDIP project[35]. A total of 22,000 patients per year will be included in EUBIROD reports.

In **Malta**, there is one central Diabetes Clinic with six peripheral clinics. A total of over 23,000 patients regularly attend Diabetes Clinics, of which 13,500 are stored on a unique computerised system. Linkage to the system for the EUBIROD reports is being organized. Currently, there are only 250 patients included in EUBIROD annual reports.

In the **Netherlands**, the West Friesland register covers nearly 5,000 diabetic subjects every year from N=75 participating primary care centres. All records are computerized through a centralized surveillance system. Patients receive care by the diabetes care system in addition to the care delivered by their own GP using a centrally organized database that is available to all involved caregivers[36]. Diabetes nurses and dieticians perform an annual follow-up examination of individual patients to assess glucose control, cardiovascular disease risk profile, and the presence of complications, and to coordinate care among different healthcare providers, including GPs, specialists, and podotherapists. Nearly 5,000 subjects followed up by the West Friesland register on a regular basis will be included in EUBIROD reports.

In **Norway**, the Norwegian diabetes register for adults is owned and financed by the National Health Board Bergen, with daily operation run by NOKLUS. The mandate is to develop a national quality register for adults by collecting data from general practices and hospitals and to give feedback reports to the participating centres. The EUBIROD annual reports will include 2,800 patients from the Bergen region.

In **Poland**, the Silesian Diabetic Centre in Katowice covers approximately 850 children and adolescents (0-18 y) with Type 1 Diabetes in a region of nearly 4,000,000 inhabitants. Work from the centre has allowed to conduct research as part of the EURODIAB program, during which 1,385 new cases of diabetes mellitus type 1 were recognized, leading to accurate estimates of the standardized incidence rate across the region[37]. A total of 800 subjects will be included in EUBIROD diabetes reports.

In **Romania,** the Institute Paulescu collects data relative to nearly 6,000 patients from the Bucharest region. Electronic databases have been developed based on the Diabcare dataset since 1990, in collaboration with Telemedica Consulting. SincroDiab, a synchronized diabetes register for the routine clinical practice in a LAN and for the design of long-term trials and epidemiological studies based on the GEHR (Good European Health Record) architecture was developed and tested in this framework[38]. Nearly 2,900 active patients will contribute to annual EUBIROD reports.

In **Scotland**, the Tayside diabetes register refers to a region located on the East coast of Scotland with a population of 394,000 inhabitants. Clinical information referring to over 17,000 diabetic patients has been ascertained electronically via the Diabetes Audit and Research in Tayside Study (DARTS) Web System, coordinated by the University of Dundee since 1999[39]. Nearly 19,000 active patients per year from the region will be reported in EUBIROD.

In **Sweden**, the Skaraborg Primary Care Database (SPCD) was initiated in the year 2000 by linking information from the 24 public health care centres (HCCs) in the county of Skaraborg in Sweden. SPCD was one of the first large databases of this kind launched in Sweden, including data on diabetes patients that have been routinely used for research[40]. A total of over 11,000 active patients per year is followed up by the system. The data from the SPCD can be linked to external databases, such as registers from Statistics Sweden and the Swedish Prescribed Drug Register,

after permission from the Central Ethical Review Board. Nearly 11,000 active patients per year will be included in EUBIROD reports.

In **Slovenia**, all children at the onset of Type 1 diabetes are admitted to the University Children Hospital, University Medical Centre Lubljana, where they are further monitored in the outpatient clinic[41]. A total of 1,600 patients per year will contribute to the EUBIROD database.

2.2 PIA Questionnaire

Scope

A PIA questionnaire[42] has been used to acquire detailed information on how data is processed by centres affiliated to the EUBIROD consortium.

Scope of the questionnaire is:

- to determine the level of privacy protection of any registry/database to be linked in the EUBIROD information system
- to evaluate how heterogeneous is the implementation of privacy principles/requirements among participating centres
- to identify key areas of concern in the implementation of privacy principles/requirements across participating centres

Contents

The content of the questionnaire is based on privacy principles enshrined in International data protection legislation, which herein have been defined as key elements of data protection (factors). As an initial step of the EUBIROD privacy impact assessment, the legal expert has reviewed the relevant privacy literature, already analysed in the BIRO project, in order to ascertain which privacy principles/norms were involved in data processing operations occurring in EUBIROD registers.

The key elements of data protection (factors) used for the analysis of the processing operations occurring in the management of diabetes registries are as follows:

- *Accountability of Personal Information,* which relates to issues such as the custody and control of personal information, third parties involvement, etc.
- *Collection of Personal Information*, relative to the authority to collect, the necessity of the information collected (minimality principle), the use of information for secondary purposes, the provision of anonymization for planning, management and/or evaluation purposes
- *Consent,* on the necessity to gather informed consent for the collection and processing of data in the registry and on how consent is obtained, if it is clear and unambiguous, if the capacity to give consent has been taken into account
- *Use of Personal Information,* focusing on the authority to use information, the application of the purpose specification principle, the use of personal identifiers for data linkage
- *Disclosure and Disposition of Personal Information*, relating to the consent/authority to disclose personal information, to the disclosure of personal identifiers, etc.
- *Accuracy of Personal Information*, dealing with the possibility for individuals to access, assess, discuss or dispute the accuracy of his/her record
- *Safeguarding Personal Information*, related to security measures and processes
- *Openness*, with regard to the provision of communication processes and to the way personal information is managed/protected

- *Individual Access to Personal Information*, which evaluates the practical implementation of access rights
- *Challenging Compliance*, which investigates the availability of complaint procedures and mechanisms to ensure accountability
- *Anonymisation Process for Secondary Uses of Health Data,* which analyse the whole compliance with international technical standards and principles

As a result, the EUBIROD PIA questionnaire is composed of N=11 sections (factors), each containing a series of questions over the same topic. The questionnaire analyses each data processing operation against privacy principles enshrined in EU and International legislation. A total of N=57 questions have been used to populate sections of the questionnaire.

The content of the questionnaire is integrally reported in Appendix 1.

A variable number of questions (sub-factors) is present in each section: N=5 for the accountability of personal information, N=10 for collection of personal information, N=6 for consent, N=5 for the use of personal information, N=5 for disclosure and disposition of personal information, N=6 for accuracy of personal information, N=8 for safeguarding personal information, N=2 for openness, N=4 for individual access to personal information, N=3 for challenging compliance and N=3 for the anonymisation process for secondary uses of health data (N=3).

For each question, partners have been asked to provide “yes” or “no” responses. Further options included "N/D" (not determined), for situations in which the register management is still at an early stage, or "N/A" (not applicable) where the specific question did not apply to the specific context. A "Provide Details" column has been used to explain specifically how a particular requirement is met or why it is not met, and whether it has been used to provide specific authoritative references. The optional field offered the opportunity to include additional comments to facilitate interpretation and resolution of any potential misunderstanding. Partners have been recommended to provide comments, details and discussion points in accurate and comprehensive manner. "Discussion Points" related to the questions have been placed at the end of each section.

Administration and data collection

The final version of the questionnaire has been distributed to all partners at the first BIRO technical meeting (Rome, November 2009) as an empty Word document to be filled in remotely (with the collaboration of local database administrators, register managers, and the task leader) and sent electronically to the Coordinating Centre

Completed data has been requested for the next meeting, scheduled after six months. During this timeframe, continuous legal advise has been provided to ensure the correctness and completeness of answers and to avoid any potential misunderstanding in the interpretation of the questions. The identity of the register submitting the questionnaire could not be blinded to both the legal expert and the statistical analyst, due to the continuous feedback required to improve the quality and the correct interpretation of the data collected.

Various rounds of submissions/corrections have been undertaken to complete the process and allocate all answers correctly. Partners have finalized the questionnaire soon after the Special BIRO Academy Meeting, held in Rome in June 2010.

2.3 Privacy Factors and the Scoring System

Sections in the EUBIROD PIA questionnaire refer to specific “privacy factors” (e.g. collection of personal information) that relate to specific EU and/or international data protection principle or norm to assess the level of compliance of diabetes registries for selected areas of interest. Each factor envisages a level of privacy protection on a specific privacy issue.

Factors provide summary results that are easy to interpret for all questions included in the questionnaire. Each section is composed of questions that can be seen as “sub-factors” (e.g. are secondary uses contemplated for the information collected?) drilling down into specific procedures that are relevant to fulfill privacy goals.

To deliver a quantitative analysis for all questions and factors included in the questionnaire, standardized coding mechanisms have been applied.

The scoring system adopted takes into account the adherence to privacy principles or norms for all identified processing operation occurring in EUBIROD centres.

The original values saved in the excel sheet (0 = No, 1 = Yes, 2 = Not Applicable / Not Determined, 9 = Missing - Blank or Comment Only) have been recoded to provide balanced sums across each section.

As a first step, all missing values have been considered equal to 0 (no compliance to privacy). All questions have been then recoded by assigning a mark of one to any privacy protective conduct, not necessarily corresponding to a response of “yes”, as interpreted by the legal expert.

Finally, factors have been computed as the sum of recoded values of the original responses.

The algorithms applied for all questions and associated factors are reported in detail in Tables 2.1-2.11. As it can be noted, not all questions entered the factor algorithm for each factor.

Scaled factors for each register have been computed as a percentage of the factor score on the total attainable score.

The overall level of privacy protection has been computed as a composite indicator obtained as the average of all scaled factors for each participating register.

Table 2.1. Accountability of Personal Information

QUESTIONS	YES	NO	NA/ND
1.1	1	0	0
1.2	1	0	0
1.3 + 1.4 If 1.3 YES and 1.4 NO/NA Otherwise	 0 1		

Factor A1 = 1.1 + 1.2 + (1.3 + 1.4) [Range: 0-3]

Table 2.2. Collection of Personal Information

QUESTIONS	YES	NO	NA/ND
2.1	1	0	0
2.2	1	0	0
2.3	1	0	0
2.4	1	0	0
2.5 + 2.6 + 2.7 If 2.6 NO or 2.7 NO Otherwise	 0 1		
2.8	1	0	0

Factor A2=2.1+2.2+2.3+2.4+(2.5+2.6+2.7)+2.8 [Range: 0-6]

Table 2.3. Consent

QUESTIONS	YES	NO	NA/ND
3.1	1	0	1
3.2	1	1	0
3.3	1	0	1
3.4	1	0	1
3.5	1	0	1
3.6	1	0	1

Factor A3=3.1+3.2+3.3+3.4+3.5+3.6 [Range: 0-6]

Table 2.4. Use of Personal Information

QUESTIONS	YES	NO	NA/ND
4.1	1	0	0
4.2	1	0	0
4.3	0	1	1
4.4	1	0	1

Factor A4=4.1+4.2+4.3+4.4 [Range: 0-4]

Table 2.5. Disclosure and Disposition of Personal Information

QUESTIONS	YES	NO	NA/ND
5.1	1	0	1
5.2	1	0	1
5.3	0	1	1
5.4	0	1	0
5.5	1	0	0

Factor A5=5.1+5.2+5.3+5.4+5.5 [Range: 0-5]

Table 2.6. Accuracy of Personal Information

QUESTIONS	YES	NO	NA/ND
6.1	1	0	0
6.2	1	0	1
6.3	1	0	1
6.4	1	0	1
6.5	1	0	1
6.6	1	0	0

Factor A6=6.1+6.2+6.3+6.4+6.5+6.6 [Range: 0-6]

Table 2.7. Sageguarding Personal Information

QUESTIONS	YES	NO	NA/ND
7.1	1	0	0
7.2	1	0	0
7.3	1	0	0
7.4	1	0	0
7.5	1	0	0
7.6	1	0	0
7.7	1	0	0
7.8	1	0	0

Factor A7=7.1+7.2+7.3+7.4+7.5+7.6+7.7+7.8 [Range: 0-8]

Table 2.8. Openess

QUESTIONS	YES	NO	NA/ND
8.1	1	0	1
8.2	1	0	1

Factor A8=8.1+8.2 [Range: 0-2]

Table 2.9. Individual Access to Personal Information

QUESTIONS	YES	NO	NA/ND
9.1	1	0	1
9.2	1	0	1
9.3	1	0	1
9.4	1	0	1

Factor A9=9.1+9.2+9.3+9.4 [Range:0-4]

Table 2.10. Challenging Compliance

QUESTIONS	YES	NO	NA/ND
10.1	1	0	1
10.2	1	0	1

Factor A10=10.1+10.2 [Range: 0-2]

Table 2.11. Anonymisation

QUESTIONS	YES	NO	NA/ND
11.1	1	0	1
11.2	1	0	1
11.3	1	0	1

Factor A11=11.1+11.2+11.3 [Range: 0-3]

2.4 Statistical Analysis

Statistical analysis included descriptive frequencies of questions, factors and overall scores with the associated 95% confidence intervals. A variety of exploratory graphs have been produced to display results obtained for the PIA questionnaire relative to single questions, factors, and the overall questionnaire.

Absolute frequencies obtained for single questions/factors using the original coding structure have been displayed using **histograms**. Results show individual values and sums obtained using the original values recorded in the Excel data sheet and saved as a comma delimited text file. In all cases, the horizontal axis indicates the response obtained, while the vertical axis the absolute number of registers for which a specific level has been recorded.

In the case of factors, histograms indicate the absolute scores obtained by the sample of registers, which may vary among different factors according to the number of sub-factors contained. For instance, if a given factor is composed of six sub-factors (questions), the maximum attainable score would be six. The vertical axis shows how many partners attained a particular score for the individual factors. The same histograms were reproduced to show the distribution of overall scores expressed as the average percentage achieved for all factors by each register.

Results for individual factors rescaled as a percentage of the maximum achievable have been summarized through a **table** including the code of the factor, its description, the number of questions contributing to the factor (determining its maximum achievable score), and measures of centrality/dispersion including the arithmetic mean, the standard deviation, the median and the range in terms of average percentage obtained. The 95% confidence intervals were computed assuming a normal distribution. Lower, upper limits defined by the interval: mean±1.96*(standard deviation).

Boxplots have been used to graphically compare the distribution of standardized factors.

The privacy profile of individual registers (in anonymous format) has been graphically displayed using **starplots**. In this figure, each starplot represents a separate register with eleven rays, where each one represents a factor ordered clockwise. To facilitate comparison across registers, the scale of all rays is fixed between 0 (dot) an 100% (predetermined maximum length for all starplots). The different shapes highlight different patterns of privacy implementation: the larger the figure, the better is the performance. A legend including factor codes is displayed at a side to facilitate the interpretation of the figure.

A **heat map** has been produced as a graphical representation of average values of a variable obtained for all factors in a two-dimensional map. Colors have been scaled around levels of yellow-red in descending order (red: lower levels of privacy). In this figure, association trees (dendrograms) displayed on each side show the result of a cluster analysis indicating the similarity between factors and registers. Each element is progressively linked to the most similar one until the whole sample is classified as one group. By selecting a stopping rule based on the sequential association, it is possible to identify separate groups according to the average value of the target variable.

Dotplots have been produced anonymously for each register to display the average position for each factor and the overall score, compared to the average and confidence intervals for the entire sample.

Ad hoc software written in the R statistical language[43] has been specifically developed to produce all the above analysis. All the source code produced to deliver the results contained in the present report is included in Appendix 2.

2.5 IT Platform

The continuous application of the EUBIROD privacy impact assessment methodology would allow a regular update of the results obtained by the initial survey through the direct interaction with end users.

To this end, an online interface has been specifically developed, defined as the "privacy performance self-evaluation platform".

The web platform includes an electronic version of the questionnaire and a management system that allows its applicability to new users with subsequent validation of submitted questionnaires by the task leader. Validated files are automatically translated into coded data submitted to R routines that update results on the server. Graphical outputs summarize the scores achieved by each centre and allow comparing own profile against the entire sample. All outputs are provided in anonymous format, so that no centre is able to identify results relative to other centres.

Technical details on the online platform are described as follows.

The electronic format is a website from where authorized users are able to fill in the questionnaire and submit it to the EUBIROD Coordination Centre located in Perugia (Italy) for revision, validation and analysis. The web platform is directly available through the BIRO Academy homepage (http://www.eubirod.eu/academy) at the following address: http://questionnaire.eubirod.eu.

The platform includes a web questionnaire (Figure 2.1) entirely reproducing the hardcopy version. A total of N=11 sessions/pages are included in the electronic form. Each page includes all questions appearing in the original form. Definitions are available to users on the "how to complete the questionnaire" session. Sessions 3 and 11, considering the relevance of their content, also include a "help" button that allows browsing information on the meaning of "informed consent" and "anonymisation" according to the latest ethical guidelines produced by the European Commission.

Functional requirements of the web-questionnaire are hereafter described:

Filling the questionnaire

- The user has the possibility of filling in the questionnaire by accessing a dedicated Web Page. In order to fill in the questionnaire, the user will be authenticated. Only Authenticated users with the right permissions will be able to fill in and complete the questionnaire.

Filling single questions

- The user does not need to fill in all the questionnaire sections in a particular order or in a specific assigned time. Each section has a "Save" button allowing the user to persist data into the database and recall them at any time. The database is matched against a User Unique Identifier or "Primary Key".

Available answers

- The user will be presented with multiple choices for each question; mapped values are 0 (no); 1 (yes); 2 (ND/NA); 9 (Missing)

Clearing available answers

- Each section of the questionnaire is provided with an optional button to "Clear Answers for this Section". If the user clicks on this button, all the answers for that particular section will be deleted and, therefore, classified as "Missing" (valued "9").

Questionnaire validation

- The IT platform envisages the possibility for an authenticated user, with the role of "PIA Validator", to validate and eventually amend answers entered by each single user. This

functionality is behind a secured area within the website, which constitutes an added layer of security for this area. The PIA Validator can edit or delete questions and can delete or clear answers given by a particular user. The difference between deleting and clearing answers is that by deleting answers, the PIA Validator physically may remove records from the database. By clearing all the answers, the PIA Validator may classify answers as "missing", with a value of 9. Furthermore, the PIA Validator is responsible for promptly informing the user of any change applied to the record and to seek confirmation to the answers given.

- Currently, these changes can be tracked through the storage of all CSV files (automatically generated at any questionnaire submission), which are automatically sent via e-mail to both the Coordination Centre and the PIA validator.
- Scope of the validation is to ensure that questions are answered correctly. For instance, the validation process may assess if questions are correctly interpreted, highlight incongruences and make sure that all sessions are answered, etc. Only validated questionnaires shall be taken in consideration for analysis.

CSV creation and submission

- Either when all sections of the questionnaire have been filled in by the user or when a PIA validator validates users' data, a CSV file is produced and sent to the Coordinator Centre in Perugia for analysis and storage. The submission is made via an email attachment and via direct download of the most recent file produced. The physical file is stored into the root folder of the web application, which is protected by server firewalls.

Non-functional requirements of the web questionnaire are hereafter described:

Programming language

- The system has been coded and implemented using Microsoft technologies, specifically:
  - ASP.NET
  - VB.NET
  - DOT NET FRAMEWORK 3.5
- The web application runs on a server machine running Microsoft Server 2003 R2 as operating System.

Back-end

- SQL Server 2008 SP1 is the Database Engine chosen.

User Interface

- The system has been provided with a user friendly interface in order to facilitate access to users with limited computing experience users and/or low frequency of usage.
- It is expected that a very short training (ten minutes) should be sufficient for properly using the system.

The conceptual model for the entire system is graphically described in Figures 2.2-2.4.

As a final step, the CSV file produced by the on-line system can be directly submitted to the R source code running the statistical analysis and producing the graphics. In a future release, the results delivered by the R program can be directly included in the user interface and provided back to the users automatically.

Figure 2.1: Web Questionnaire

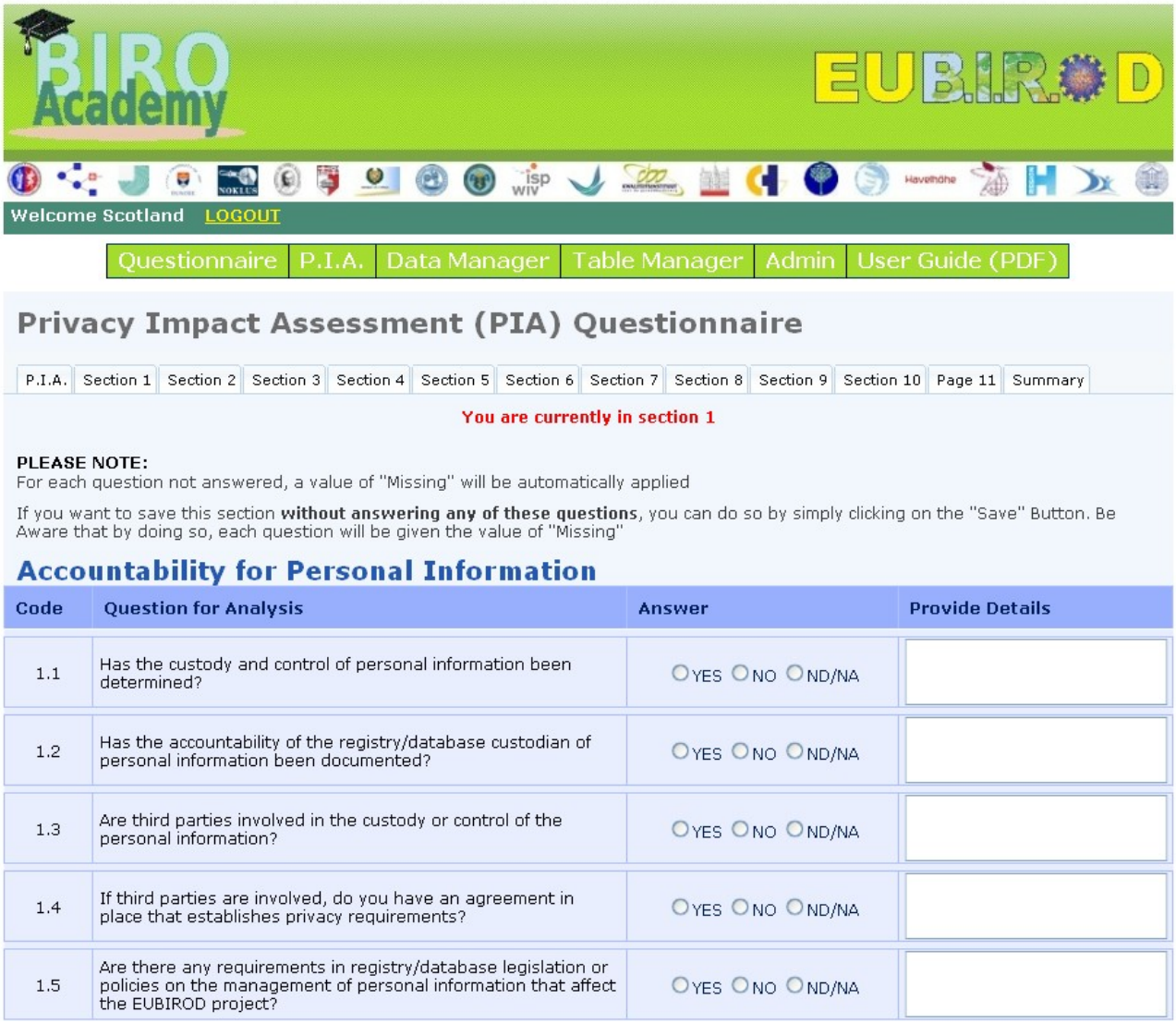
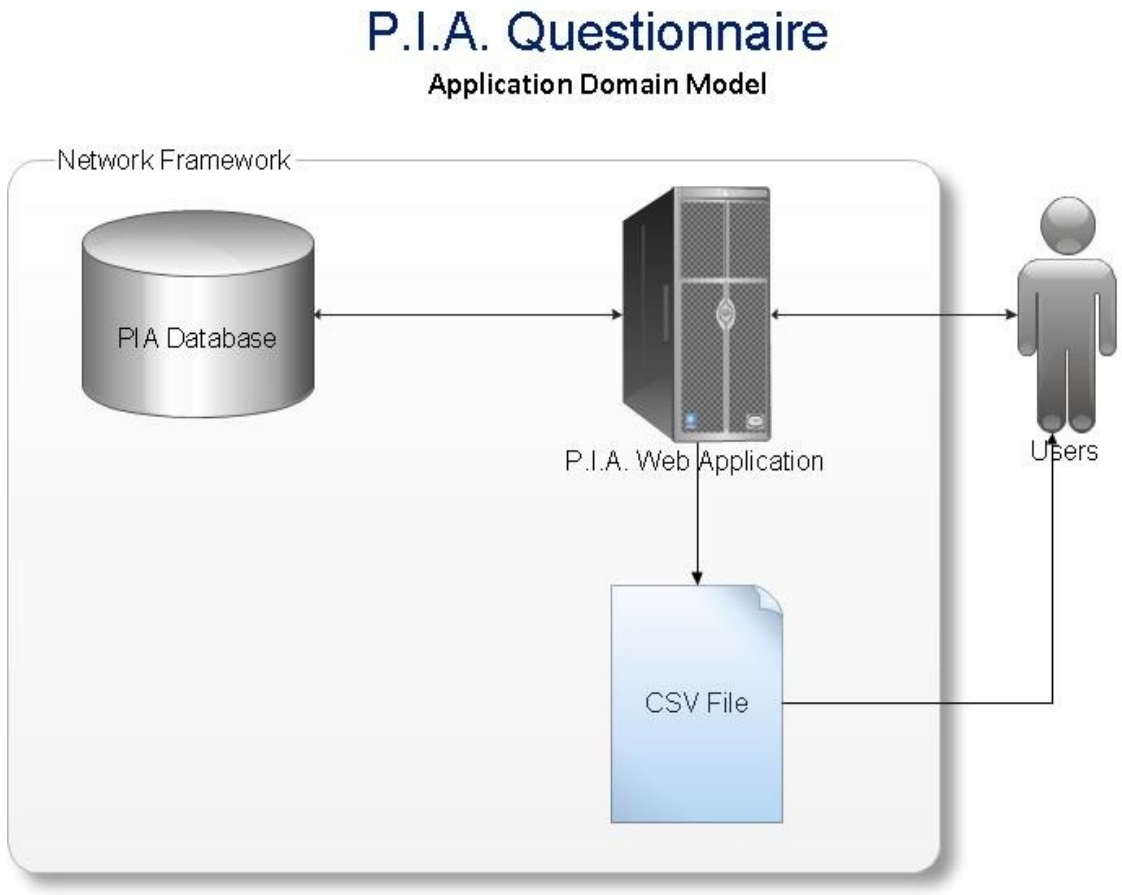


Figure 2.2: Application Domain Model



Legend:

1. PIA database: The Back end Database used to store and retrieve data.
2. User: The authenticated Users and Pia Validator who will be using the system.
3. CSV FILE: The File produced by the web application, stored into the Web Application root folder and sent to the User and Pia Validator via Email attachment.

Figure 2.3: Conceptual Website Diagram

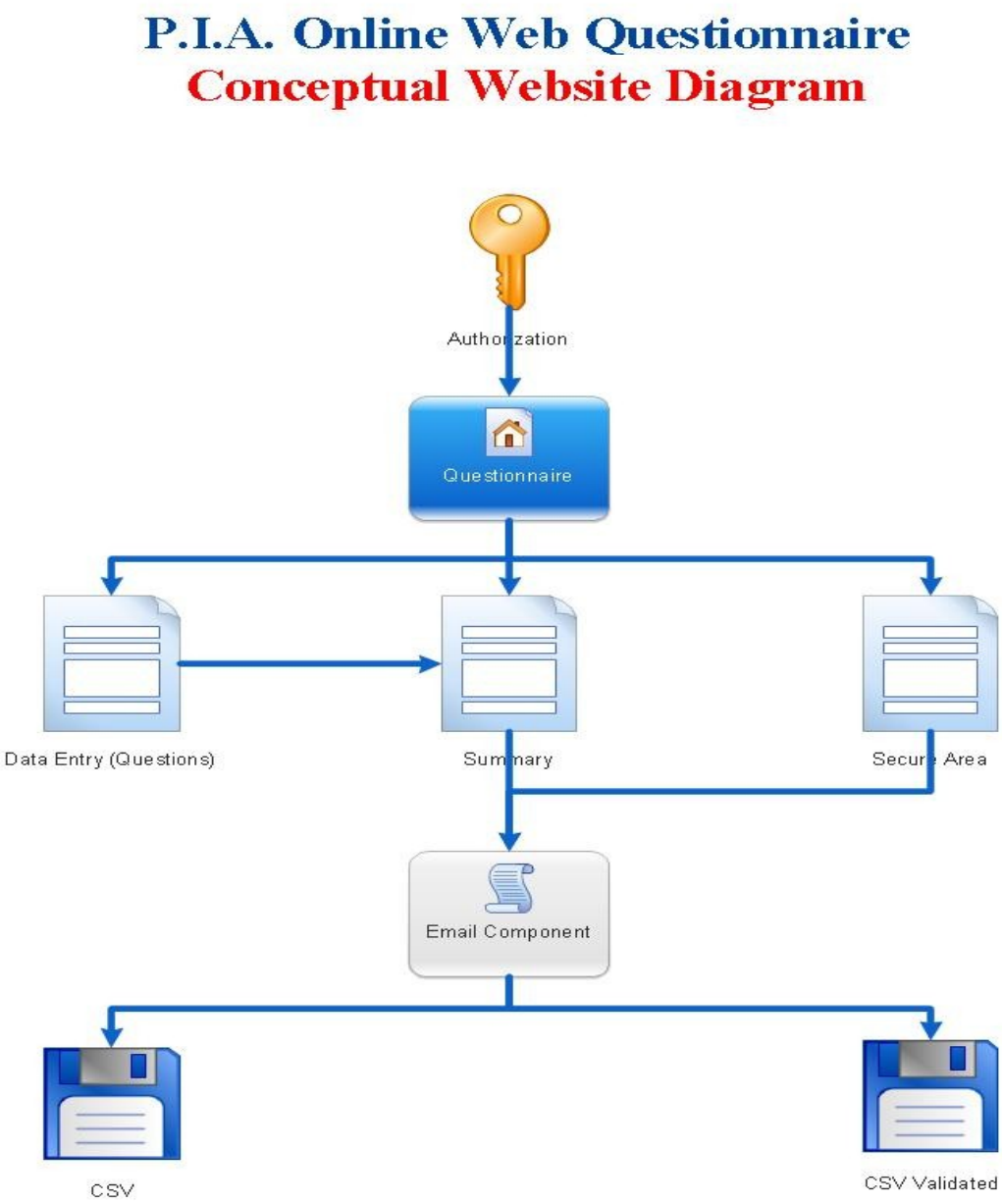
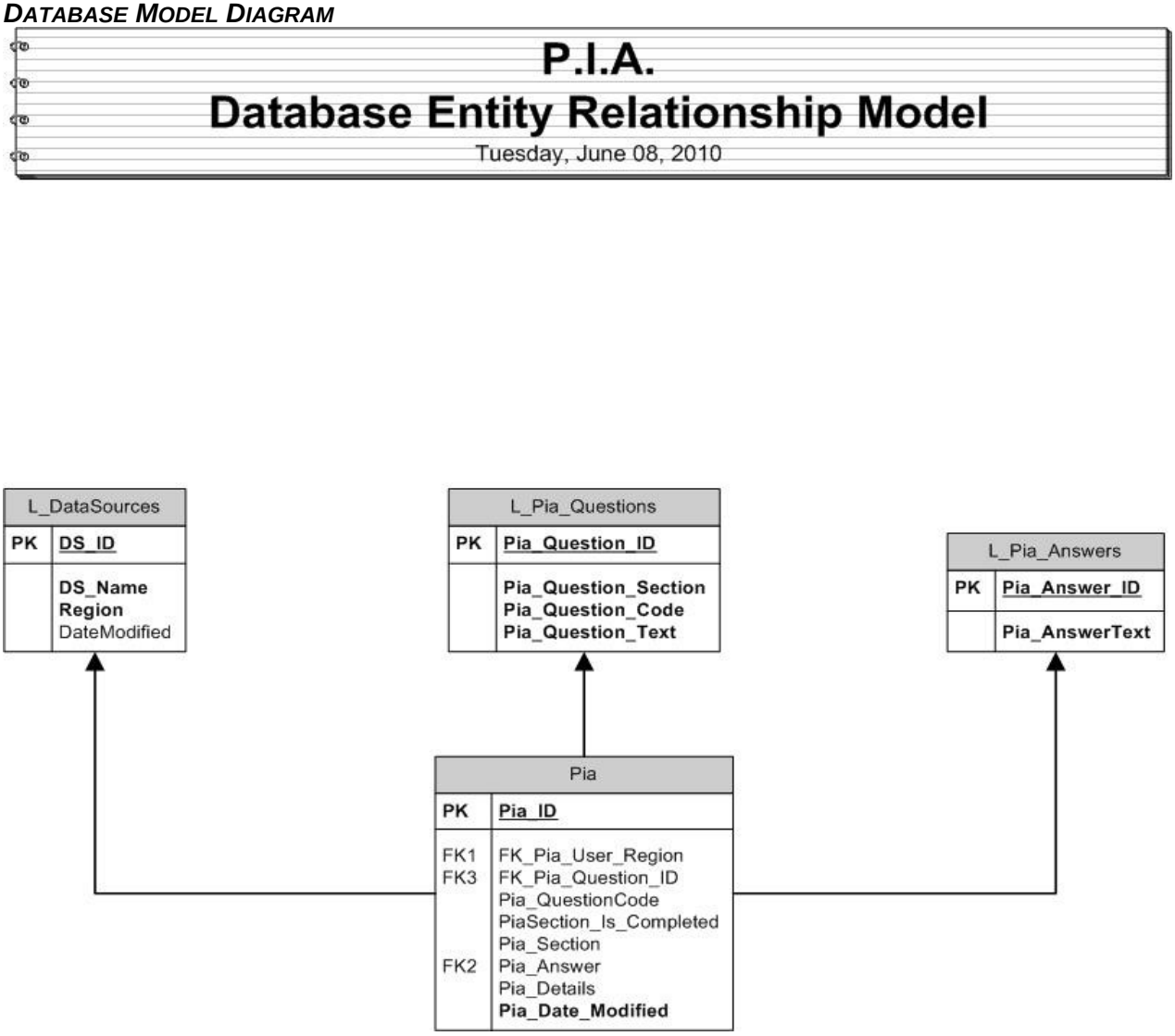


Figure 2.4: Database Model Diagram

DATABASE MODEL DIAGRAM



3. Results

Results relative to the N=18 respondents to the EUBIROD privacy impact assessment questionnaire are hereafter presented into four sections:

1. Main findings from single questions
2. Factors, providing an overview of the procedures applied by responding centres in terms of privacy protection for any factor identified
3. Overall privacy performance evaluation, reporting the overall level of privacy protection achieved by all centres
4. Privacy performance self-evaluation, presenting the visual display that will be provided to each register to evaluate own performance against the average observed for the whole sample and the highest attainable level.

3.1 Main Findings from Single Questions

The eleven sections contained in the PIA questionnaire describe a broad range of elements related to the respect of privacy legislation that should be duly taken into account in the management of diabetes registers. The results obtained by the entire set of N=49 questions are extensive and can be overwhelming for the amount of details provided. To this end, sections would be more efficiently summarized by the presentation of factor results.

This section of the report focuses on the results obtained for specific questions that indicate the degree of heterogeneity in the implementation of privacy protective procedures across Europe. The following sections are presented in detail: collection of personal information, consent, use of personal information and the anonymisation process for secondary uses of health data.

Section 2. Collection of Personal Information

Responses to the PIA questionnaire have shown that the Collection of Personal Information is performed directly from the individual in N=12 (67%) cases, while in N=6 (33%) not directly from the individual (Figure 3.1).

The purpose for which personal information has been collected is documented in N=13 (76%) registries, while in the remainder 24% is not documented (Figure 3.2).

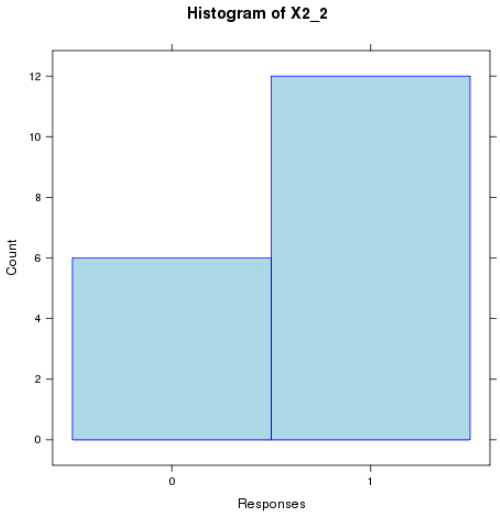


Figure 3.1: Responses to Question 2.2

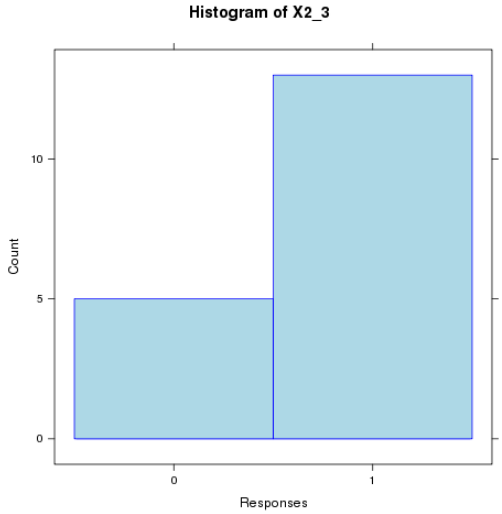


Figure 3.2: Responses to Question 2.3

A total of N=17 (94%) cases collect information that is necessary for the registry according to the minimality principle (Figure 3.3).

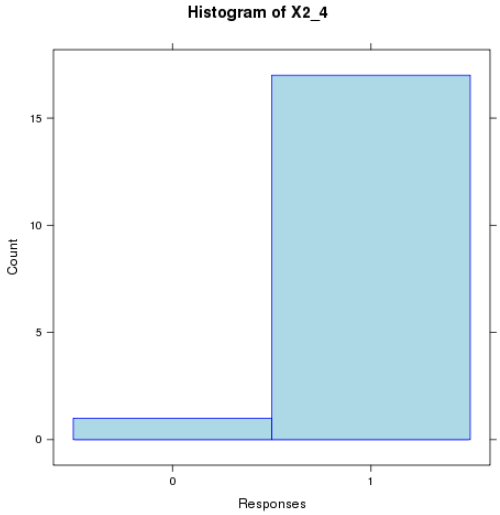


Figure 3.3: Responses to Question 2.4

The possibility of using data contained in the registry for secondary purposes is envisaged only by N=8 (44%) registers (Figure 3.4).

However, if data is to be used for purposes not previously identified, informed consent is required in N=12 (67%) cases and not required in N=4 (22%) cases (Figure 3.5).

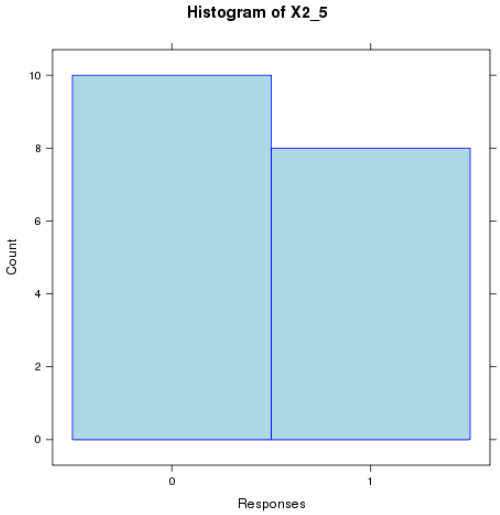


Figure 3.4: Responses to Question 2.5

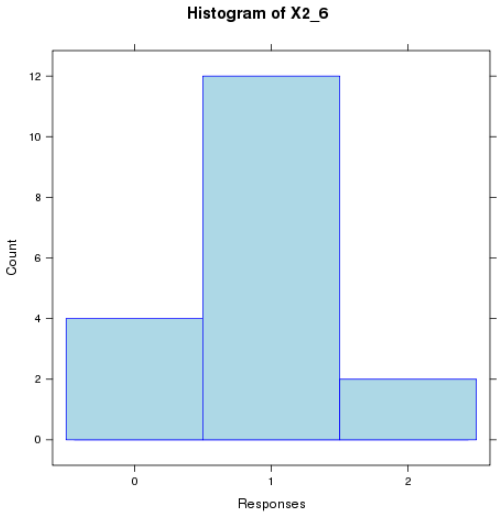


Figure 3.5: Responses to Question 2.6

When consent is not required, as highlighted by responses to question 2.7, consent requirements are waived by previously acquired authority to use and disclose personal information; e.g.: authorized by law (Figure 3.6).

A total of N=17 (94%) cases has reported that anonymisation procedures are applied for planning, management and/or evaluation purposes (Figure 3.7).

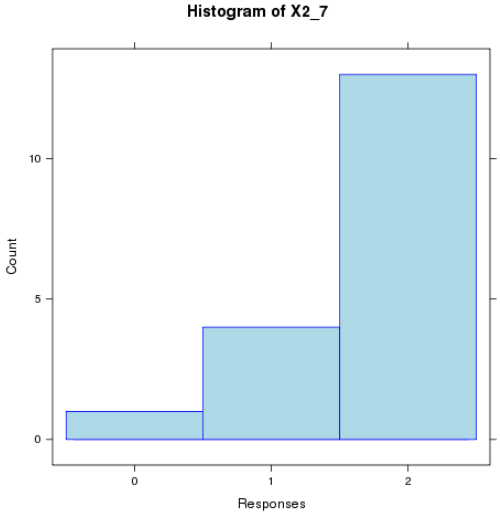


Figure 3.6: Responses to Question 2.7

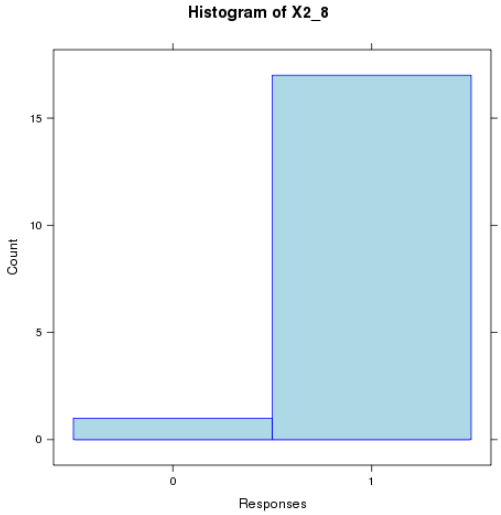


Figure 3.7: Responses to Question 2.8

The possibility to collect some personal information from public databases is envisaged only in N=4 (22%) registries, while N=13 (72%) registries do not have access to them (Figure 3.8).

The collection from multiple sources through a common patient identifier is performed by N=6 (33%) registries, while the remainder 67% do not have access to multiple sources (Figure 3.9).

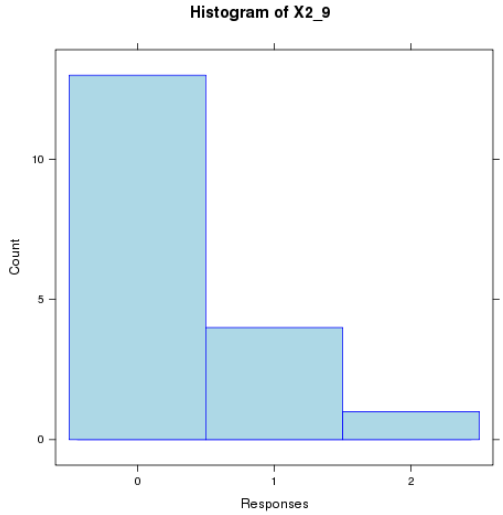


Figure 3.8: Responses to Question 2.9

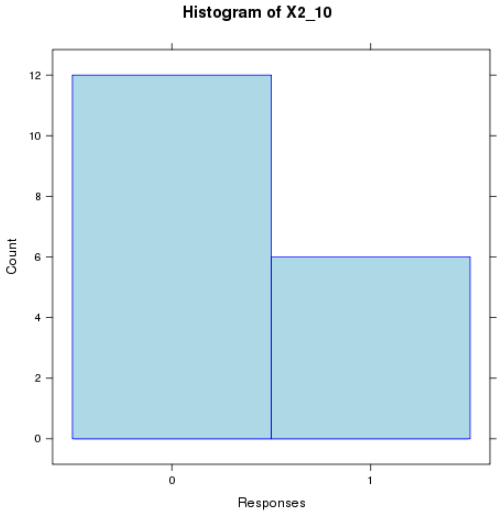


Figure 3.9: Responses to Question 2.10

Section 3. Consent

Consent is required by N=11 (61%) centres in order to collect and process data in the registry, while is not required in N=7 cases (Figure 3.10). Accordingly, in all cases where consent is necessary, it is obtained directly from the individual; and not directly from the individual in all the other cases (Figure 3.11).

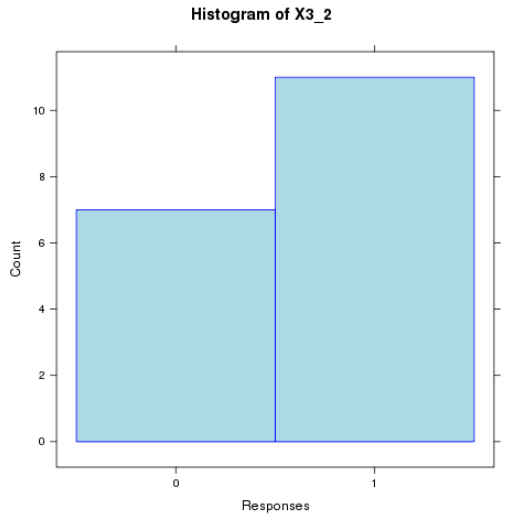


Figure 3.10: Responses to Question 3.2

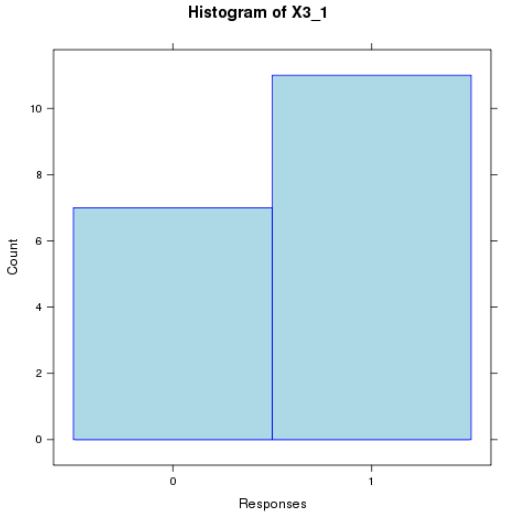


Figure 3.11: Responses to Question 3.1

Section 4. Use of Personal Information

Data linkage is performed by N=9 (50%) registries, while N=8 (44%) centres do not link across multiple databases (Figure 3.12). Data matching is consistent with the stated purposes for which personal information has been collected in N=11 (61%) cases (Figure 3.13).

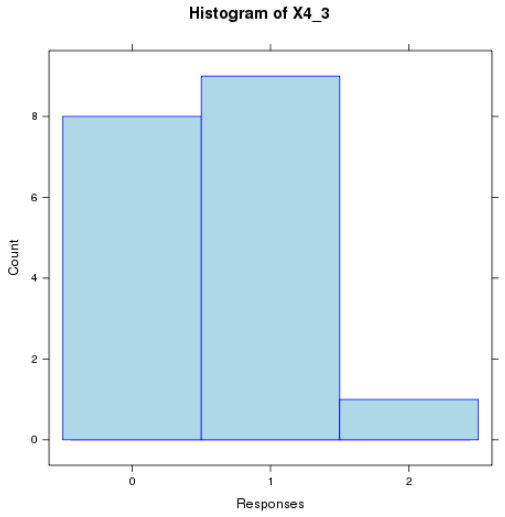


Figure 3.12: Responses to Question 4.3

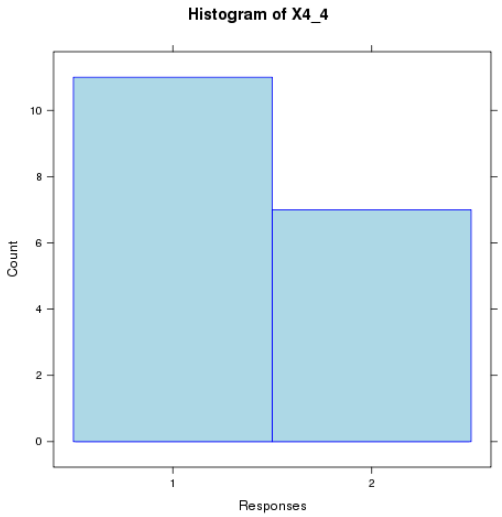


Figure 3.13: Responses to Question 4.4

A notification to the Privacy Commissioner is required in N=7 (39%) cases (Figure 3.14).

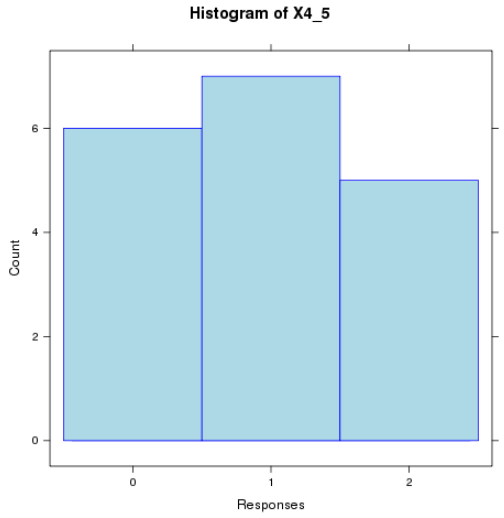


Figure 3.14: Responses to Question 4.5

Section 11. Anonymisation Process for Secondary Uses of Health Data

Standard anonymisation procedures are envisaged in N=11 (61%) registers before further processing data for secondary uses, while they are not performed in N=4 (22%) cases (Figure 3.15).

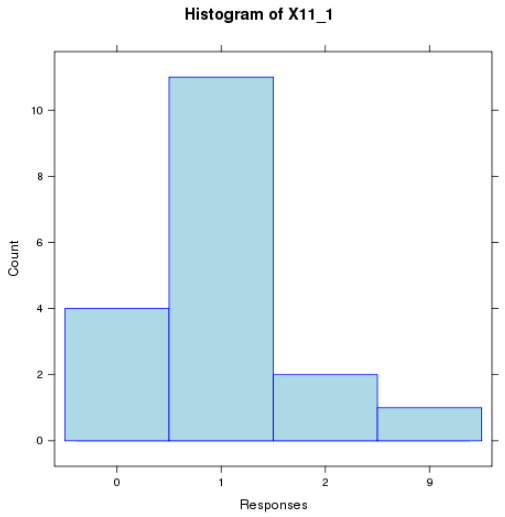


Figure 3.15: Responses to Question 11.1

Where applied, standard anonymisation procedures are also:

- in compliance with international technical standards and continuously updated according to the state of the art (Figure 3.16)
- in compliance with the Data Protection Principles; for instance, performed confidentially, providing information to patients about the processing operation, applying security mechanisms for data storage and retention, etc. (Figure 3.17)

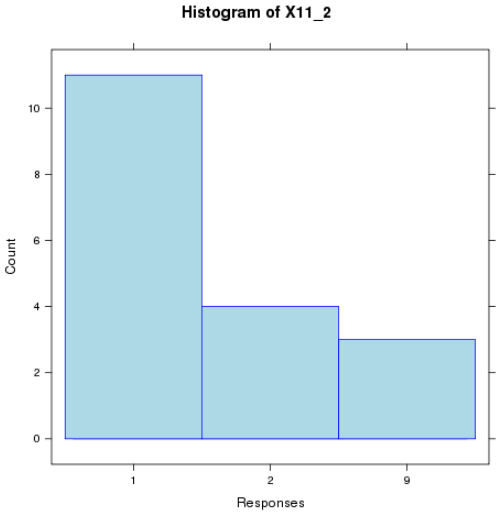


Figure 3.16: Responses to Question 11.2

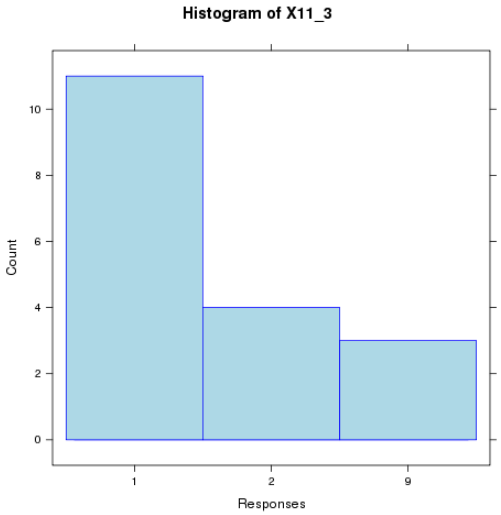


Figure 3.17: Responses to Question 11.3

3.2 Factors

This section reports in detail the results obtained for all key privacy factors identified by the EUBIROD privacy questionnaire. The detailed presentation of the absolute values obtained as a sum of the individual components (questions, or sub-factors) is followed by a statistical summary and a graphical display of standardized values, expressed as a percentage of the maximum achievable for each factor.

Section 1. Accountability of personal information

This section of the questionnaire refers to the the custody and control of personal information.

Questions included in this section have been selected to assess:

- if the custody and control of personal information are determined and documented
- whether there is any involvement of third parties

Responses show that N=17 (94%) registries/databases have the custody and control of personal information both determined and documented (Figure 3.18).

The involvement of third parties is envisaged only in N=4 (22%) registries/databases; however, in those cases the involvement of third parties is disciplined by an agreement that establishes privacy requirements.

Results for this factor are fairly homogeneous. Indeed, the highest score (MaxS=3) was reached by N=17 (94%) registries/databases, with only N=1 (6%) register recording a low score of one.

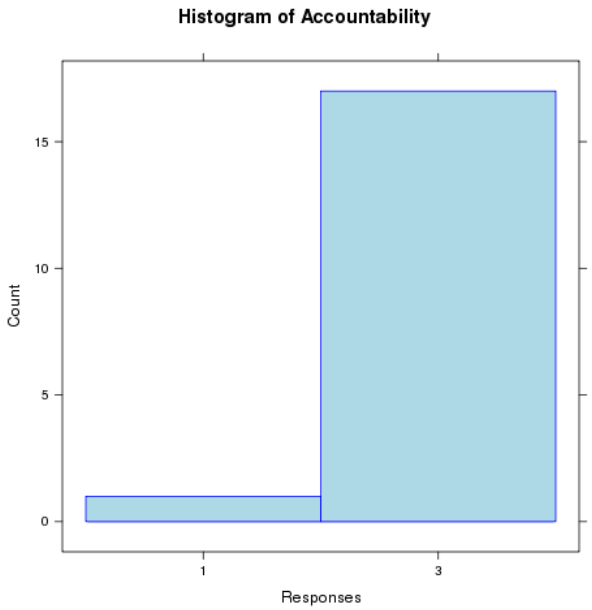


Figure 3.18: Acccountability of Personal Information

Section 2. Collection of Personal Information

This section deals with the collection of personal information, aimed to assess:

- the authority to collect
- if information is collected directly from the individual
- the necessity of the information collected (minimality principle)
- if secondary uses are contemplated
- if anonymization is performed when information is used for planning, management and/or evaluation purposes

Results from the questionnaire (Figure 3.19) show that values are concentrated around near optimal levels. A total of N=6 (33%) registers reached the maximum score (MaxS=6), N=7 (39%) centres had a score of five, N=4 (22%) a score of four and N=1 (5%) a score of three.

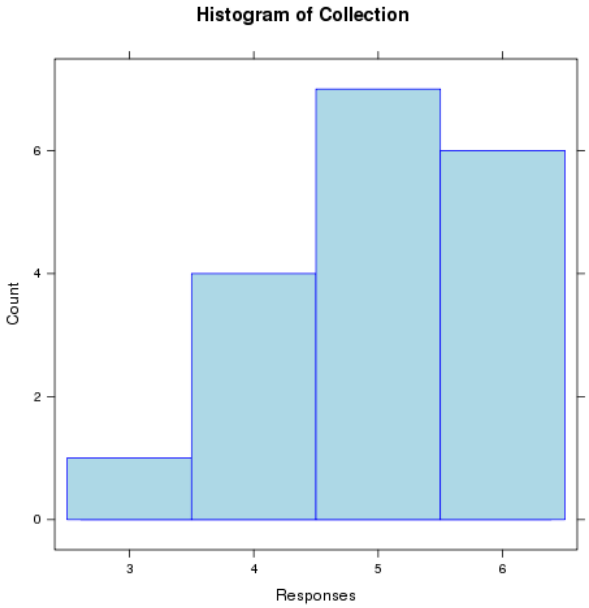


Figure 3.19: Collection of Personal Information

Section 3. Consent

This section explores informed consent issues in order to determine:

- if consent is obtained directly from the individual
- how consent is obtained
- if consent is clear and unambiguous
- if consent is needed for the collection and processing of information in the registry/database
- if the capacity to give consent is taken into account

The high heterogeneity of scores obtained by EUBIROD centres for the “consent” factor is clearly shown in Figure 3.20. One third of the centres gained a maximum score (MaxS=6), the remainder covering the whole range: N=3 (17%) obtained a score of five, N=4 (22%) a score of four, N=2 (11%) a score of three, N=1 (6%) a score of two and N=2 (11%) a score of one.

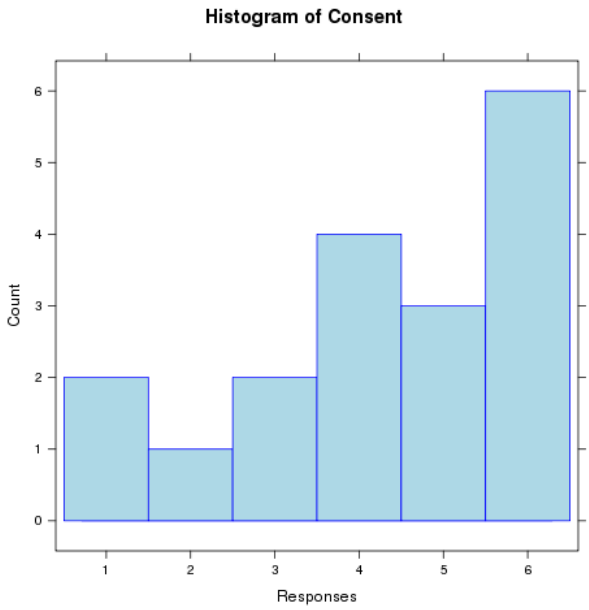


Figure 3.20: Consent

Section 4. Use of Personal Information

This section is aimed at analysing how information is used within the EUBIROD centres. Questions are intended to evaluate:

- the authority to use information
- the purpose specification principle
- the use of personal identifiers for data linkage

Results for this section were fairly homogeneous, as shown in Figure 3.21. Although only N=1 (6%) centre obtained the maximum score (MaxS=4), most centres still reported a high score, with N=16 (78%) scoring three and N=1 (6%) a value of one.

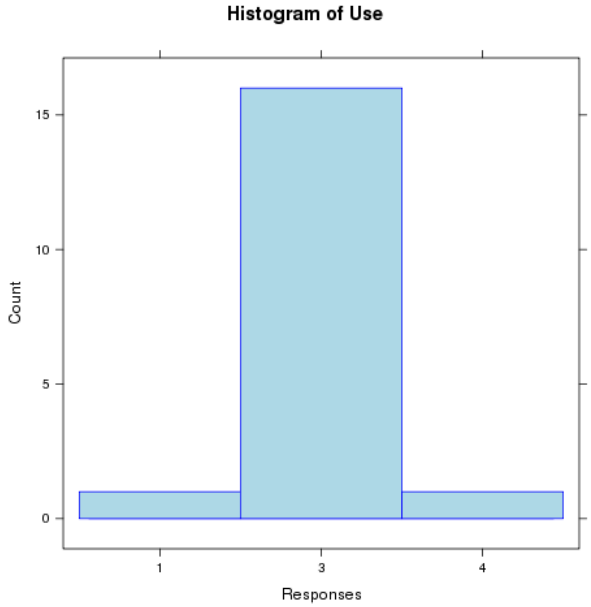


Figure 3.21: Use of Personal Information

Section 5. Disclosure and Disposition of Personal Information

This section refers to specific issues surrounding the disclosure and disposition of personal information.

Questions are set up to assess:

- if consent is required to disclose personal information
- the authority to disclose without consent
- if personal identifiers are disclosed
- if transborder data flow is performed
- if disposition of personal information is required

Figure 3.22 highlights that none of the centres reached the maximum score (MaxS=5). The highest score obtained was three for N=5 (28%) centres, while the bulk of N=12 (66%) recorded a score of two and N=1 (6%) a score of one.

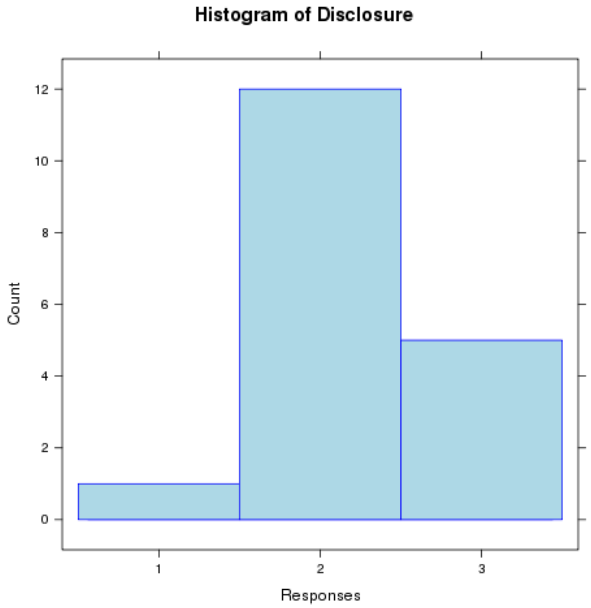


Figure 3.22: Disclosure and Disposition of Personal Information

Section 6. Accuracy of Personal Information

This section investigates the accuracy of personal information and the possibility for individuals to access, assess, discuss or dispute the accuracy of his/her record.

To this end, questions concern:

- the existence of standard procedures to ensure that personal information is accurate, complete and up-to-date
- if record is kept of: a) changes occurred; b) requests for review of errors or omissions; c) corrections; d) any decision not to correct
- if notice of corrections made to health records is given to the data subject
- if a set procedure allows the individual to access, assess and dispute the accuracy of his/her data

Figure 3.23 shows the heterogeneous results obtained for this factor. A total of N=4 (22%) centres obtained the maximum score (MaxS=6), N=5 (28%) a score of five, N=3 (17%) a score of four, N=3 (17%) a score of three, N=2 (11%) a score of two and N=1 (5%) a score of one.

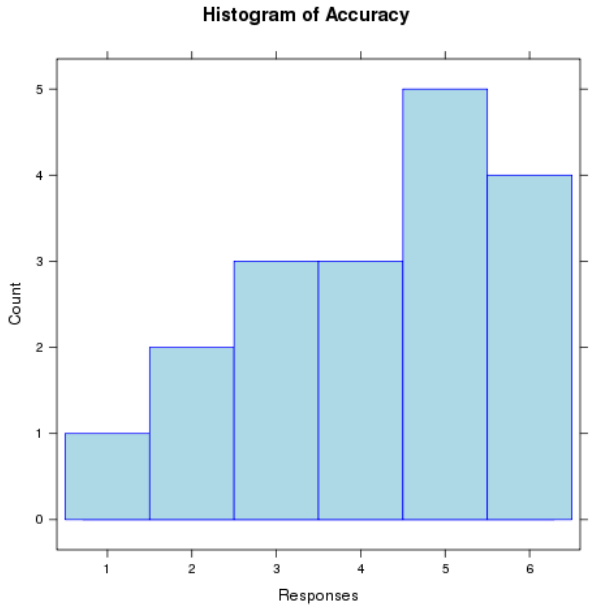


Figure 3.23: Accuracy of Personal Information

Section 7. Safeguarding Personal Information

This section is concerned with security measures for safeguarding personal information. Questions aim to ascertain:

- if security procedures are documented
- personnel training on security
- if and how security controls are put in place
- if security measures applied are commensurate to the sensitivity of information
- if contingency plans for security breaches are envisaged and documented
- if security measures are subject to to quality assurance audit

Figure 3.24 shows the distribution of values obtained for this factor. The maximum score (MaxS=8) was reached by one third of the centres. The others covered the whole spectrum of possibilities: N=3 (17%) presented a score of seven, N=5 (28%) a score of six, N=2 (11%) a score of 5, N=1 (6%) a score of four and N=1 (6%) a score of three.

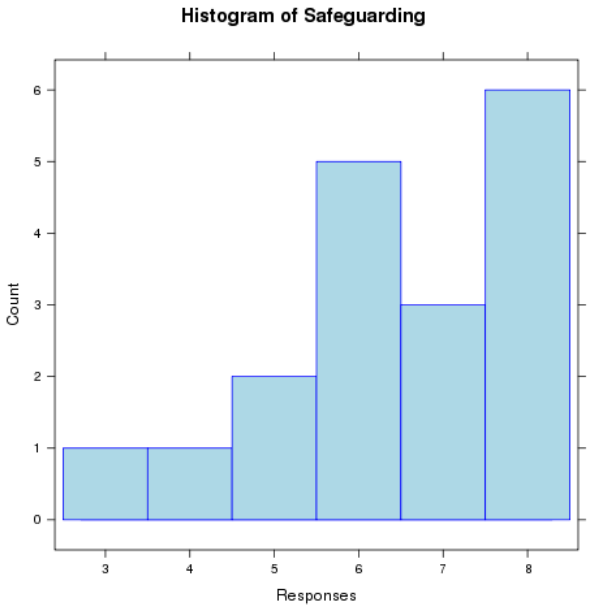


Figure 3.24: Safeguarding Personal Information

Section 8. Openess

This section has been structured to assess openess to the public of personal information managed and protected within the centre.

To this end, questions relate to:

- the existence of a communication plan
- the existence of a predetermined process that allows individuals to easily access such information

Figure 3.25 provides a sketch of the results obtained for this factor, which were mostly concentrated close to the optimal level. A total of N=12 (67%) registries obtained the maximum score achievable (MaxS=2), N=5 (28%) a score of one and N=1 (5%) a score of zero.

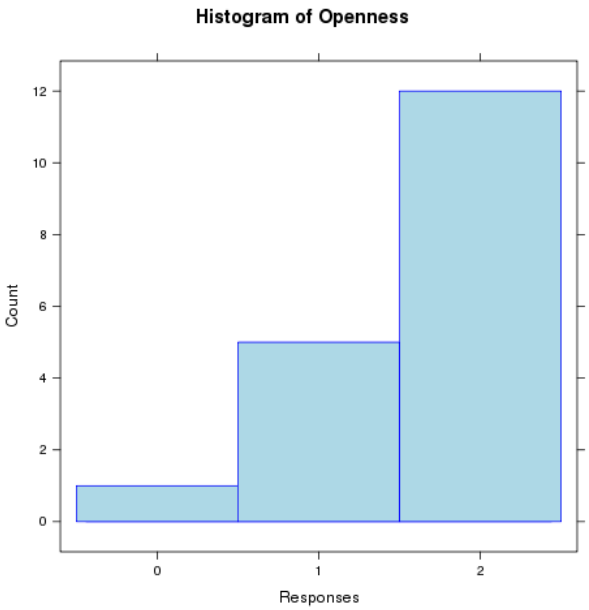


Figure 3.25: Openness

Section 9. Individual Access to Personal Information

This section relates to access rights.

Specific questions aim to assess:

- if the system is designed to allow individual's access to own personal information
- if eventual corrections are notified
- if custodians are aware of access rights
- if “routine” access is envisaged

Results shown in Figure 3.26 indicate that individual responses were normally distributed around a central value of two. The maximum score (MaxS=4) was obtained by N=2 (11%) centres, a score of three by N=4 (22%) centres, a score of two by N=9 (50%) centres, a score of one by N=2 (11%) centres and a score of zero by N=1 (6%) centre.

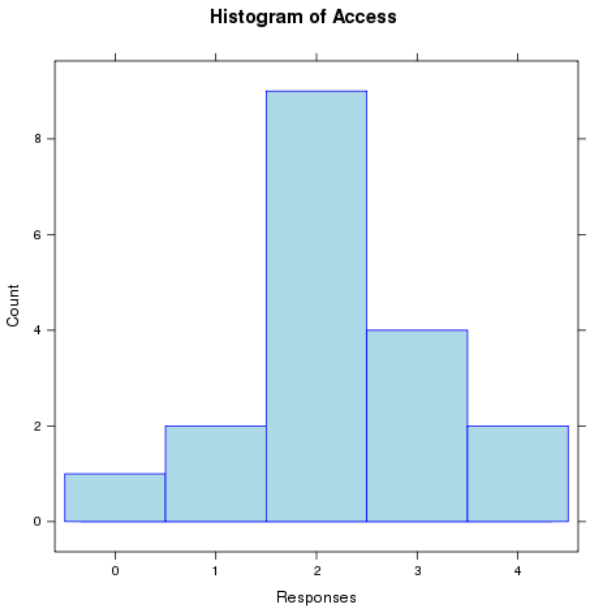


Figure 3.26: Individual Access to Personal Information

Section 10. Challenging Compliance

This section explores issues surrounding the availability of complaint procedures and mechanisms to ensure accountability.

Figure 3.27 shows results obtained for this factor. The maximum score (MaxS=2) was attained by N=12 (67%) centres, with N=3 (17%) centres scoring one and N=3 (17%) centres zero.

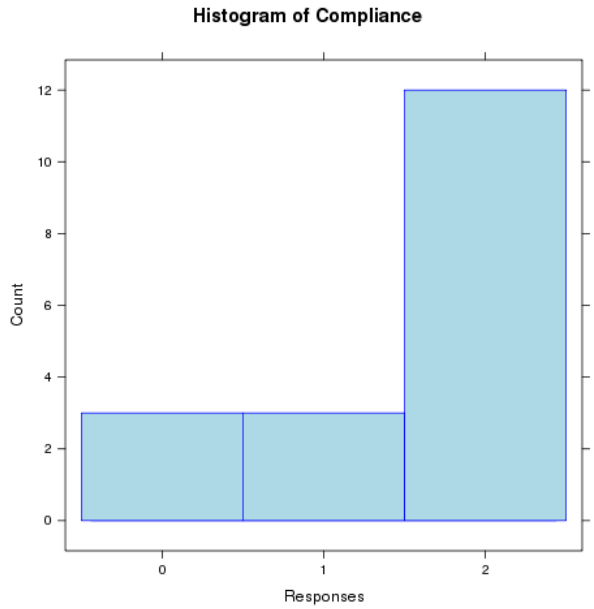


Figure 3.27: Challenging Compliance

Section 11. Anonymisation Process for Secondary Uses of Health Data

This section assesses the anonymisation process to ascertain:

- if a standard procedure is envisaged
- if it is compliant with international technical standards
- if individual data, before anonymisation, are processed according to privacy requirements

Results are shown in Figure 3.28. The distribution of this factor was fairly heterogeneous. Although the maximum score (MaxS=3) was reached by N=12 (67%) registries, a score of two was present in N=3 (17%) registries, a score of one by N=1 (6%) centre and a score of zero by N=2 (11%) centres.

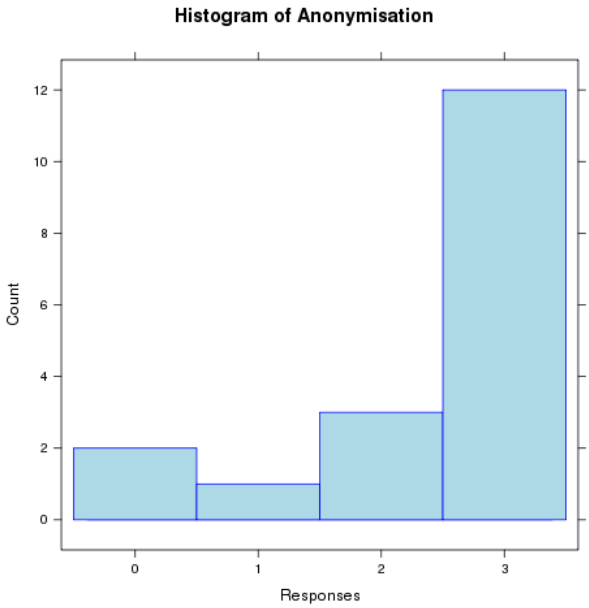


Figure 3.28: Anonymisation Process for Secondary Uses of Health Data

Standardized comparison of factor results

A standardized comparison of factor results, including statistical measures for each factor and the overall average as a percentage of the maximum attainable is presented in Table 3.1.

Median values show that the following areas should be regarded as the most problematic:

- Disclosure and Disposition (40%)
- Individual Access (50%)

The following factors, presenting a median equal to 75%, are also of concern:

- Consent
- Use of Personal Information
- Accuracy

Over 50% of the registers included in the EUBIROD sample recorded a maximum score (100%) for the following factors:

- Accountability
- Openness
- Anonymisation
- Challenging Compliance

Average results should also be compared to their variability, expressed by the standard deviation and the range of variation. These measures may reveal gaps in the implementation of privacy principles that can have negative consequences on the comparability of information across Europe and can be difficult to resolve at a late stage of planning.

For instance, in the case of anonymisation and compliance, a median at 100% may obscure the fact that the arithmetic mean, potentially prone to outweigh outlying values, is indeed much lower (79% and 75% respectively). In fact, the standard deviation (average deviation from the mean) is equal to 35% and 39% respectively, which can be considered fairly high. In the case of compliance, a total of N=3 (17%) registers scored zero.

In the EUBIROD sample, the following factors showed a high variability of scores (standard deviation, range):

- Challenging Compliance (39%, 0-100%)
- Anonymisation (35%, 45-100%)
- Openness (30%, 0-100%)
- Consent (28%, 17-100%)
- Accuracy (26%, 17-100%)
- Individual Access (25%, 0-100%)

By the way, results show that most factors are not normally distributed, meaning that in this case the mean and standard deviation may not represent adequate measures of centrality/dispersion. In each case, both the mean/standard deviation and median/range must be jointly examined. It should be also noted that factors with a highest number of categories have a higher likelihood of showing a higher dispersion of values.

Figure 3.29 summarizes the above results using boxplots to compare the distribution of all factors for all registers participating to the EUBIROD privacy impact assessment.

Table 3.1. Statistical measures for standardized factors and overall average as a percentage of the maximum attainable score

Factor	Description	No. Questions	Mean	Standard Deviation	95% C.I.	Median	Range
A1	Accountability	3	96.3	15.7	89.0-100.0	100.0	33.0-100.0
A2	Collection	6	83.3	15.1	76.3-90.3	83.3	50.0-100.0
A3	Consent	6	71.3	28.5	58.1-84.4	75.0	16.7-100.0
A4	Use	4	73.6	13.5	67.4-79.8	75.0	25.0-100.0
A5	Disclosure	5	44.4	11.0	39.4-49.5	40.0	20.0-60.0
A6	Accuracy	6	69.4	25.7	57.6-81.3	75.0	16.7-100.0
A7	Safeguarding	8	80.6	18.8	71.9-89.2	81.2	37.5-100.0
A8	Openness	2	80.6	30.3	66.5-94.6	100.0	0.0-100.0
A9	Access	4	55.6	25.1	44.0-67.1	50.0	0.0-100.0
A10	Compliance	2	75.0	39.3	56.8-93.2	100.0	0.0-100.0
A11	Anonymisation	3	79.6	34.6	63.7-73.5	100.0	44.5-100.0
	OVERALL		**73.6**	**11.1**	**68.5-78.8**	**74.8**	**68.5-78.8**

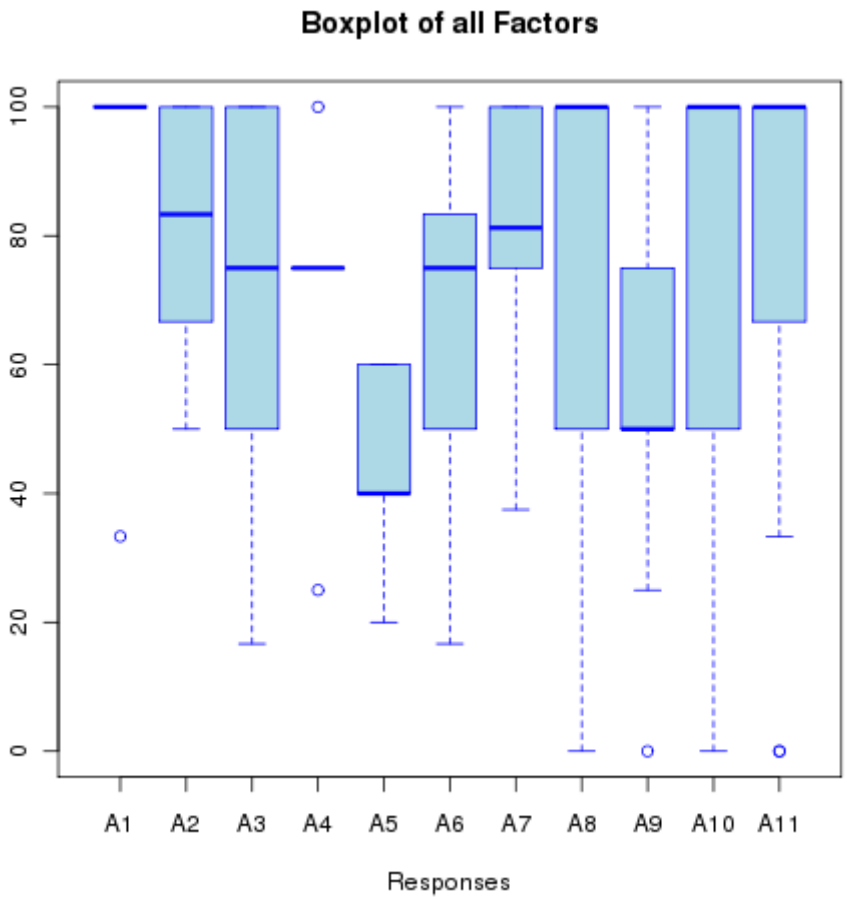


Figure 3.29: Standardized comparison of factor results

3.3 Overall Privacy Performance Evaluation

The distribution of overall scores obtained by EUBIROD registers allows evaluating the overall level of privacy performance observed against the highest attainable level of privacy protection.

Figure 3.30 shows that this distribution is approximately normal, centered around an average close to 75% of the optimal level (for details see Table 3.1). A total of N=9 registers (50%) obtained values above 75%, N=5 (28%) centres between 65%-75%, N=3 (17%) centres between 50%-65%, and N=1 (5%) below 50%.

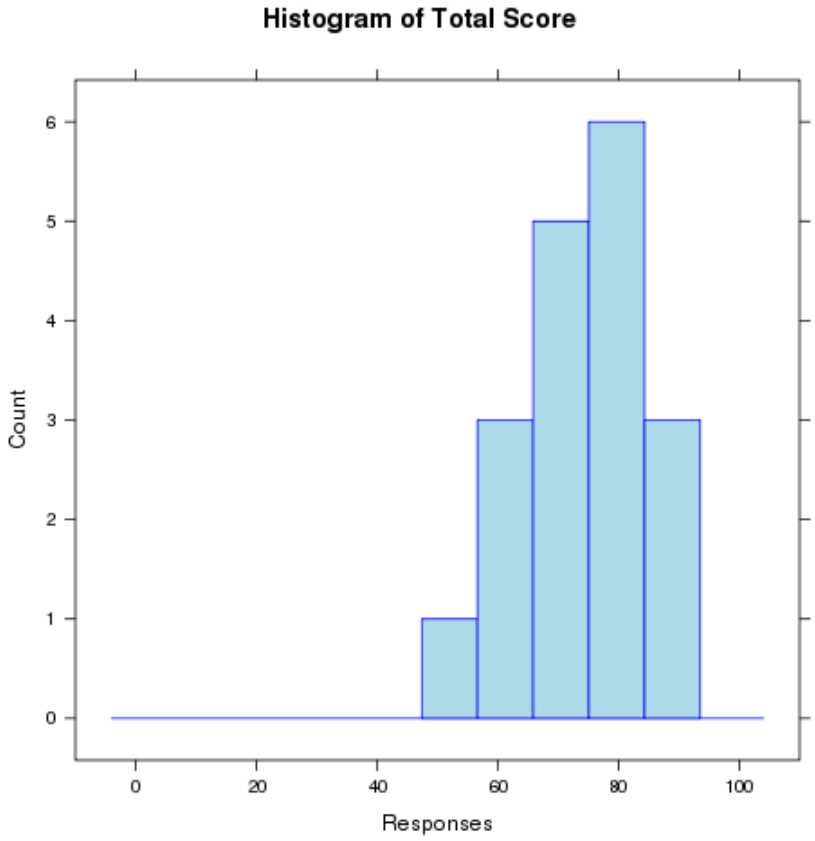


Figure 3.30: Histogram of Total Scores

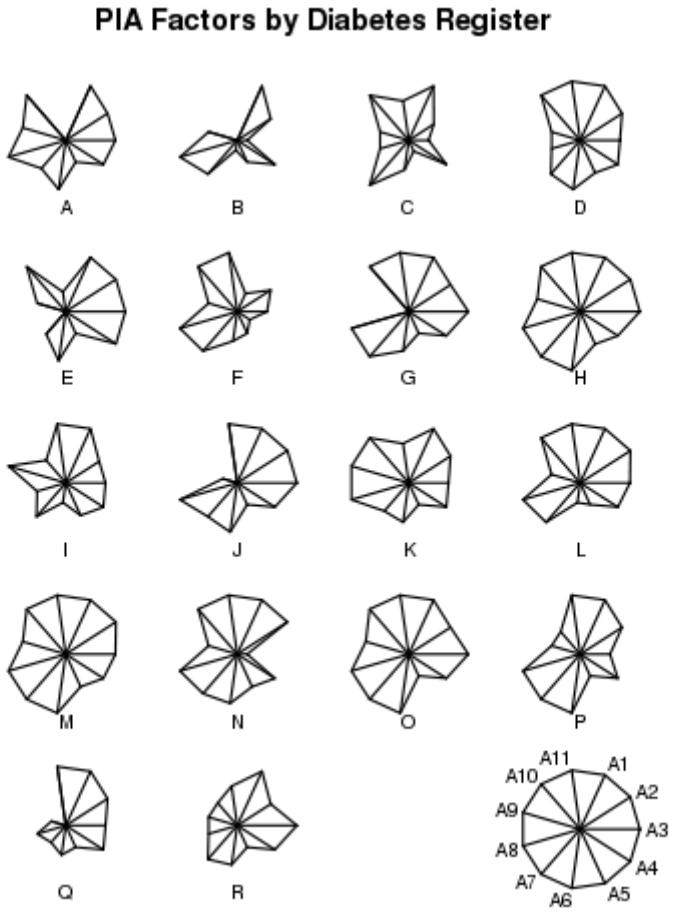


Figure 3.31: Privacy Performance of EUBIROD Registries

The “privacy profile” of each individual register can be examined using starplots (Figure 3.31). The reference figure in the lower right corner displays a legend of factor codes and represents the maximum level achievable (best profile).

At a glance, the different shapes indicate a high degree of heterogeneity across participating registers. Missing slices express deviations from the expected management of specific privacy procedures. In at least N=4 cases (labelled as B,F,Q,R) small figures indicate the simultaneous presence of low values for several factors. In N=3 cases (labelled as H,M,O) profiles are fairly close to the ideal reference.

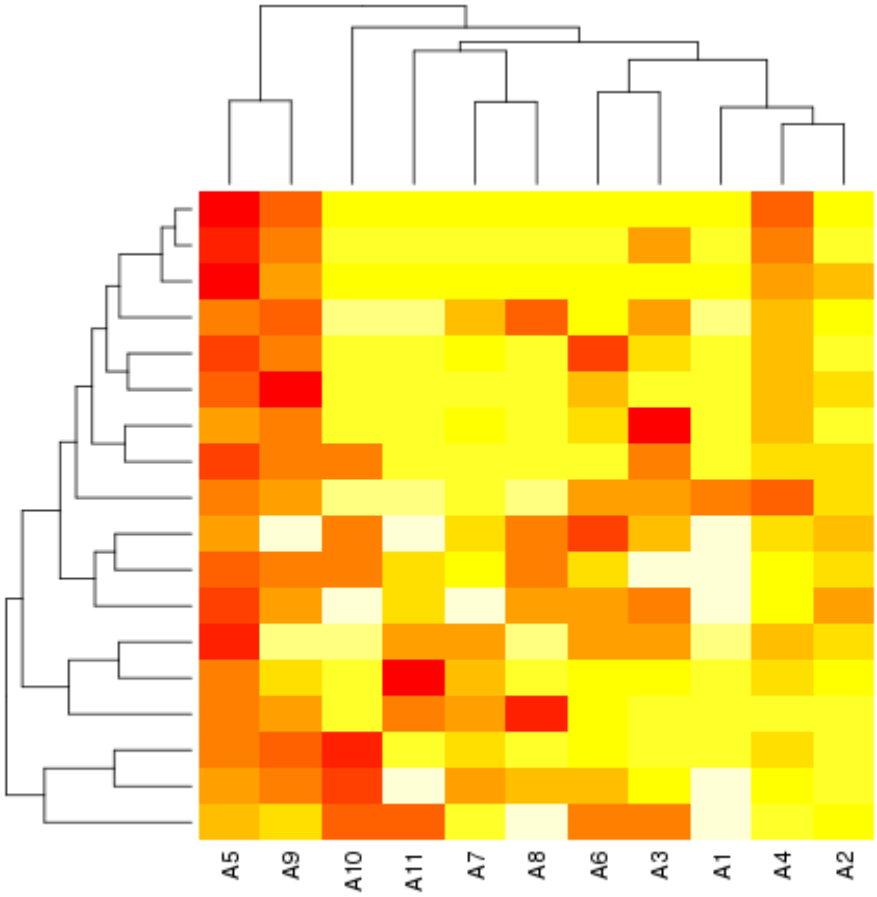


Figure 3.32: Cluster Analysis of Privacy Factors and EUBIROD Registries

The “heatmap” displayed in figure 3.32 provides a simultaneous overview of the similarities between registers and between factors, based on the average scores obtained, with the added possibility to classify them into homogeneous groups using the trees (dendrograms) displayed at a side.

Here our interest is more focused on the actual variability found across the whole sample, as the classification of factors can be very interesting to highlight the presence of key elements of concern.

The heatmap roughly identifies N=4 groups of factors, by order of performance:

- CLASS A: HIGHLY PROBLEMATIC - A5: Disclosure and A9: Access
- CLASS B: PROBLEMATIC - A10: Compliance
- CLASS C: SUB-OPTIMAL - A11: Anonymisation, A7: Safeguarding and A8: Openness
- CLASS D: NEAR OPTIMAL - A3: Consent, A1: Accountability, A4: Use and A2: Collection

These results are broadly consistent with other analyses.

However, they should only be intended as exploratory, as clustering methods are known to be unstable with small sample sizes.

3.4 Privacy Performance Self-Evaluation

A “Privacy Performance Self-Evaluation Chart” has been produced to offer the opportunity to display own results against the average to all partners.

Dotplots present the relative position of each register for each factor, along with the average of the overall sample and the associated 95% confidence intervals. The graphical display allows to identify areas of excellence at 100%, acceptability ranges, as well as privacy factors that need improvement through appropriate corrective actions.

These plots are distributed separately to each member of the privacy impact assessment team and are never published/displayed in a format unveiling the actual name of the register. This way, reports can foster collaboration and stimulate self-evaluation.

Figure 3.33 shows an example of a register anonymously labelled as “D”. In this case, the register shows a high performance compared to the bulk of the sample for almost all factors, except for “safeguarding”, “openness” and “access”, which fall outside of the confidence intervals. Optimal levels of 100% are reached for accountability, compliance and anonymization.

The “Privacy Performance Self-Evaluation Chart” represents a useful tool that can be included in the online platform to provide prompt feedback to all register managers interested in sharing their experience for the collective improvement of privacy enhancing procedures.

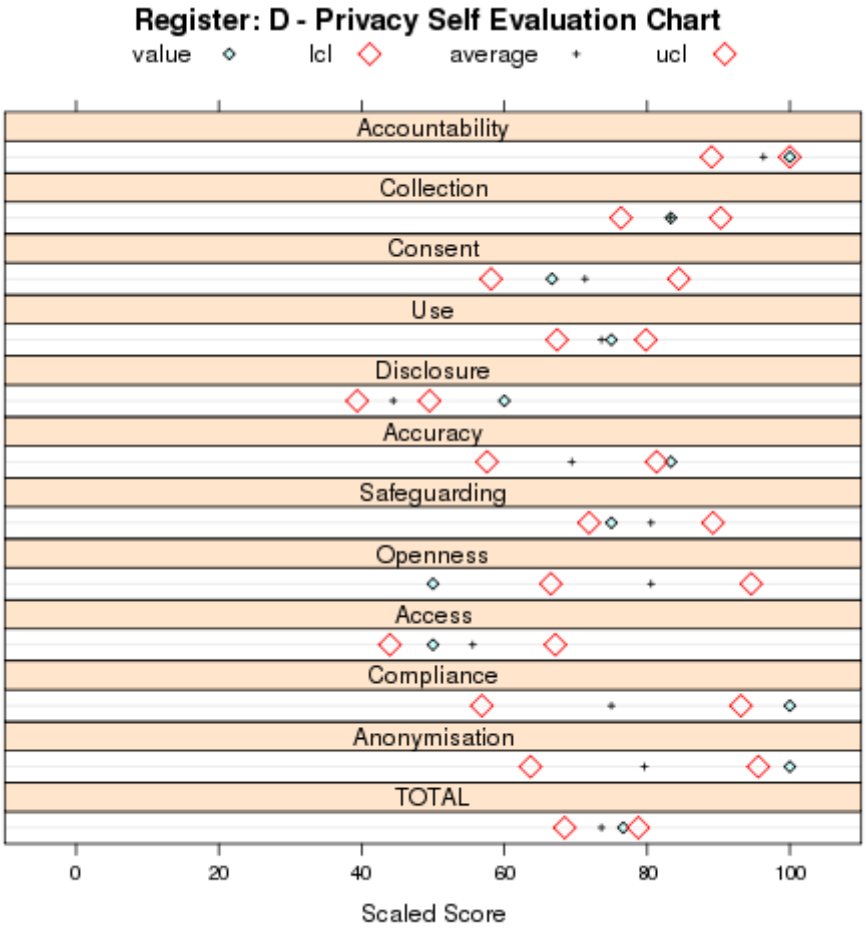


Figure 3.33: Self Evaluation Sample

4. Discussion

4.1 Research Needs

In recent years, the simultaneous evolution of health services research, statistical methods, and information technology has led to the availability of massive administrative databases and disease registries that are increasingly used to support policy[44]. Differently from "ad hoc" epidemiological studies, such sources are maintained on a routine basis and are in constant evolution. They are usually enforced by national/regional legislation for disease surveillance and for monitoring the provision of services, particularly for the control of health expenditure.

Although issues surrounding data quality are still under discussion[45], administrative data and disease registries are rapidly becoming the major information platform for sophisticated health systems research. Routine data are used in particular to produce quality and outcome indicators for performance evaluation frameworks and are ultimately needed to provide evidence-based recommendations to policy makers[46].

Access to routine databases (normally with some limitations at the governmental level, but more widely possible at the point of healthcare provision) allows to link datasets at the subject level, ordinarily without explicit patient consent, making possible to carefully control for data quality. For instance, it allows to check for double counts and to exclude from denominators those who have died or emigrated. Therefore, more precise and unbiased results can be obtained[47]. Data linkage involves access to an updated list of personal identifiers, which can lead to the identification of high spending and high-risk groups, allowing analysts to look at repeated services and to improve the precision of all estimates at population level.

Nonetheless, target information to investigate the provision of health services remains still dispersed across different classes of users and different data administrators. By definition, some of these hurdles may not be overcome, as databases are naturally gathered as a result of the provision of services taking place at different settings. Although possible, in many cases there is no framework linking those databases by automatic means, and the analyst must construct an "ad hoc" database to perform the statistical analysis. Most often, high quality clinical data contained in disease registers are not automatically integrated with other important sources e.g. pharmaceutical data, hospital discharges etc, limiting the possibility for health researchers to provide targeted recommendations for health policy.

4.2 Research needs and EU legislation on privacy protection: where is the balance?

The EU and International legislative instruments do not consider, in general terms, the right to privacy as an absolute right, but as a right that should be weighed against other matters/rights that benefit societies, including public health. Privacy norms should be therefore interpreted consistently with the goals of scientific investigation and health research, including the attainment of complete data[45]. The exemptions to the prohibition of processing operations involving personal data, e.g. those envisaged for public health, health care and health research, constitute clear examples of the non-absolute nature of the right to privacy. In other terms, privacy protection is conceived as a right that in principle should not jeopardize the right to the highest attainable level of health: health research is one of the crucial means to foster such societal right.

In line with this assumption, *Art. 1 of the Data Protection Directive*, states that principles of privacy protection should not be used to restrict the free flow of information across European countries.

The interest of societies in enhancing population health strongly depends on the possibility of conducting appropriate research in the health sector; the availability of personal data from multiple sources is fundamental to this purpose. In the case of health research, the public interest in health

monitoring at population level could be regarded as overriding the private interests of individual privacy, save that appropriate safeguards are guaranteed by legislation or by the relevant supervisory authorities.

Considering that interests of privacy protection and health research might conflict on issues surrounding the increasing demand of researchers to access data in identifiable form from different data sources, appropriate regulations should be implemented to achieve an appropriate balance between the two interests.

The EU Data Protection Directive strongly fosters the recognition and implementation of the right to privacy, but also clearly recognizes the need of societies to attain better health and health care. To this end, it provided several exemptions to the prohibition of processing sensitive data when health improvements are involved.

However, the Directive does not specify how to obtain the right balance between the two competing interests, leaving it to the implementation of the Directive at national level. For instance, the possibility for Member States to provide additional exemptions to those laid down by the Directive for reasons of public health and the development of code of conduct on the ways to anonymize data are mostly left to Member States legislation.

The actual implementation of the Directive at national level should be therefore carefully monitored to assess and understand whether, how and to what extent this balance has been achieved in practical national settings, particularly in the field of health research.

To comply with EU legislation, health systems researchers must implement data processing techniques on the ground of complex interpretations of the Directive. Different aspects must be taken into account in the study design: from informed consent to the respect of patients' rights, from data collection to use, disclosure, storage, disposition and security of data.

Research in this field would enormously benefit from a targeted action of a regulatory body proposing practical solutions to reconcile the needs and expectations of investigators with legal obligations.

The implementation of the Directive in Member States has been the focus of a survey by the Work Group on Data Protection and Confidentiality of the Health Information Strand of DG SANCO[48]. The survey has shown that the implementation of the Directive has not been consistent across Member States. Some countries have adopted national data protection legislations allowing for sophisticated information systems to process sensitive data for public health studies, health research, health monitoring, etc.

However, several Member States have made large use of the possibility to interpret the Directive more strictly by implementing more stringent privacy provisions when sensitive data are involved. In these cases, linking multiple data sources has been found either hardly possible or explicitly impeded. Therefore, it could be inferred that the balance between privacy protection and health research, envisaged by the EU Data Protection Directive, has been tipped in favour of the individual right to privacy in several Member States, producing a misinterpretation of the Directive.

The case of Estonia is emblematic in this regard. As a matter of fact, Estonia passed a data protection legislation that omits any of the exemptions (to the general prohibition of processing sensitive data) accorded by EU Directive (95/46/EC) to the processing of personal data for historical, statistical or scientific purposes[49]. As a result, the work of population-based medical registries and epidemiological research has been seriously hampered by the scarce quality and accuracy of accessible data, which has become mostly biased. Consequently, the development of evidence-based health policies and, ultimately, the improvement of public health are hardly achievable in this context. Although the EU Data Protection Directive allows for Member States to apply more stringent provisions, the Estonian legislation has certainly passed the boundaries of a

sound interpretation of the privacy principles contained in the EU legislative framework, distorting the general aims of the European legislator.

Although ethical values are already well acknowledged by the Directive, and the public interest is adequately taken into account in relevant legislation, the translation of the EU Directive into national laws has led to highly variable implementations, which in many ways may negatively impact on the efficient and effective organization of diseases registers.

Therefore, to reduce variability and to improve privacy protection in Europe, it is crucial to directly identify factors presenting the largest heterogeneity in the implementation of privacy principles and to highlight the key areas of concern in privacy protection.

The EUBIROD project has achieved the above goals through the following steps:

- an in depth review of disease registers practices
- a mixed, qualitative-quantitative approach to assess the variability in the implementation of the Data Protection Directive
- the analysis of interpretative patters
- the identification of key areas of privacy concern

4.3 EUBIROD Privacy Analysis

The survey conducted in EUBIROD through the PIA questionnaire has allowed an objective assessment of the impact of the European data protection legislation on diseases registers, in particular diabetes registers.

Scope of the EUBIROD PIA is to answer the following questions:

- How heterogeneous is the implementation of privacy requirements/principles among participating centres?
- Which are the key areas of concern on which advice and guidance is most needed?

The sample of registers included in the survey, although not representative of the state of the art across all Europe, offered a substantial overview of the topic across eighteen countries.

The PIA questionnaire has been used to collect data on all forseeble privacy issues that might be incurred in the management of diabetes registers.

The content of the questionnaire and the identification of privacy factors are based on the Canadian Privacy Impact Assessment Guidelines[42] and a review of the privacy literature. The analysis has been facilitated by the definition of a scoring system for each factor, bases on the assumption that scores for that particular issue can provide a linear measure of the level of privacy protection, according to the relevant legislation and the procedures applied in the sample of registries.

Descriptive analysis has been facilitated by recoding original questions to assign marks in terms of compliance/not compliance to privacy principles/norms.

The proposed metrics represent an initial contribution towards the realization of a fully validated system to measure the degree of heterogeneity in the implementation of privacy principles/norms and the level of privacy protection across Europe.

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

The analysis of individual factors shows that the major areas of concern (median, range) are:

- disclosure and disposition of personal information (40%, 20-60%)
- individual access to personal information (50%; 0-100%).

The following factors are also highly problematic:

- consent (75%; 17-100%),
- use of personal information (75%; 25-100%)
- accuracy of personal information (75%; 17-100%)

Factors showing on average a high variability (standard deviation) include the following:

- challenging compliance (39%)
- anonymisation (35%)
- openness (30%)
- consent (28%)
- accuracy of personal information (26%)
- individual access to personal information (25%).

The range of overall scores achieved by EUBIROD registers was 69-79% (mean:74%, standard deviation: 11%), with a median close to 75% and almost 20% of the sample falling above 80% of the maximum performance.

The EUBIROD survey has produced a detailed description of how personal information is handled in eighteen diabetes registers across Europe, allowing an identification of the key areas of privacy concern in the management of diabetes registers and an overview of the variability of approaches at European level.

Privacy performance has been measured against both absolute and mean values obtained for the whole sample.

The rationale for providing both values is that, theoretically, a perfect adherence to all privacy principles and requirements is obviously auspicable. However, providing mean values of EUBIROD Centres will allow comparing the performance of individual centres against values obtained in pratical settings. Thus, it provides fruitful information on how privacy norms/requirements have been practically implemented across Europe.

The creation of a dedicated tool to improve the management of privacy issues, herein defined as “Privacy Performance Self-Evaluation” of diabetes registers, represent an innovative tool to feed information back to individual centres. Through it, each survey respondent to the questionnaire may directly and indipendently identify own areas of concern in terms of deviation from privacy requirements and assess which areas can be improved.

The Privacy Performance Self-Evaluation tool may represent a general model of privacy performance evaluation that can also improve the quality of information contained in the registries.

For instance, the low “accuracy of personal information” found in EUBIROD may be related to to the unmet need to access/link additional data sources. Furthermore, the use of data for secondary purposes is rarely allowed. These conditions may have a negative impact on data accuracy and can ultimately hamper the research validity of the information included in diadetes registers.

The proposed methodology fosters collaboration, rather than “privacy league tables” in order to generate quality improvement loops that can increase data accuracy and completeness.

The self-evaluation tool realized in the EUBIROD project could be used as a general model of collaborative privacy performance evaluation, fostering the creation of privacy enhancing disease registers.

The findings of this survey could be used to develop targeted actions at both Europen and National levels. While the EU should provide suitable guidelines to Member States in order to foster a sound interpretation of EU legislation, Member States should ensure that individual users apply all regulations without jeopardizing health goals.

Data accuracy and completeness could be ensured by enforcing appropriate safeguards for those data processing operations that pose privacy risks. For instance, data linkage could be performed through trusted third parties that would guarantee the rispect of privacy norms.

Legislation should therefore recognize the importance of data processing operations that are crucial to improve health system performance. However, it is also fundamental that the ethical values enshrined in EU and international legislation are fully respected across Europe.

The privacy performance self evaluation tool could be used as a means to foster privacy enhancing registers and to reconcile the conflicting interests of health research and privacy.

5. Conclusions

Administrative data and disease registries are rapidly becoming popular means to deliver evidence-based information for policy.

The EU and International legislative instruments do not consider, in general terms, the right to privacy as an absolute right, but as a right that should be weighed against other matters/rights that benefit societies, including public health. Therefore, privacy norms should be interpreted consistently with the goals of scientific investigation and health research, including the attainment of complete data.

However, t**he way to obtain the right balance between the two competing interests is not paved by the Directive, being mostly left to the implementation at national level**.

The analysis performed in the EUBIROD project has confirmed that the balance between privacy protection and health research, envisaged by the EU Data Protection Directive, has been tipped in favour of the individual right to privacy in several Member States, producing a misinterpretation of the Directive.

Our investigation has directly identified, through objective metrics, the privacy principles that have been implemented heterogeneously across Europe and highlighted the key areas of concern that need targeted actions at both national and European level.

The **general model of privacy performance self-evaluation developed in EUBIROD** can help managers of disease registers to easily identify the main areas of concern, including those that can impact on the quality of information. Consequently, corrective measures could be directly implemented at the level of the individual centre.

The model fosters collaboration, rather than competition on privacy performance, in order to generate both privacy enhancing registers and **quality improvement loops** that can increase data accuracy and completeness.

A concerted action at both the legislative level and point of care provision is needed to achieve the right balance between the right to privacy and the right to the highest attainable level of health.

Appendix 1: PIA Questionnaire

How to Compile the Questionnaire

The PIA questionnaire provides a series of questions derived from EU and International privacy principles and regulations.

EUBIROD Partners should provide yes/no responses to a series of questions along with a comments section. An "N/D" (not determined) response may apply for situations where project planning is at an early stage. An "N/A" (not applicable) can be inserted where questions are not applicable.

The "Provide Details" column should be used to explain specifically how a particular requirement is met or why it is not met, or should be used to provide specific authoritative references.

"Discussion Points" related to the questions are placed at the end of each section.

If a response in the questionnaire indicates that the registry/database has no legal authority to collect, use or disseminate personal information, then immediately consult a departmental legal advisor to determine whether to proceed any further with the initiative.

The results from completing the questionnaire will be used to determine the level of privacy protection of any registry/database to be linked in the EUBIROD information system and to form the basis of the PIA Report.

Privacy is herein intended to be a broader concept than legal compliance; hence, it is recommended to provide comments, details and discussion points in accurate and comprehensive manner.

Definitions ex Data Protection Directive (95/46/EC)

PERSONAL DATA

Personal data are any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, using reasonable means, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his or her physical, physiological, mental, economic, cultural or social identity.

- Data allowing *direct identification* are data that can be easily related to a data subject and reveal their identity. This is the case of data such as the name, address, date of birth or even genetic data which, when combined with one another, allow identification with a small margin of doubt.
- *Indirect identification* requires further steps to make a link between a specific person and the data being processed. Therefore, the fact that data are not directly related to a person does not necessarily imply that they do not constitute personal data.

ANONYMOUS DATA

Anonymous data are be defined as data that cannot be qualified as personal data, since they do not (any more) allow direct or indirect identification of the data subject using reasonable means.

PERSONAL DATA PROCESSING

The concept of processing is very broad. It applies to any operation or set of operations that are performed upon personal data, whether or not by automatic means. Data processing is considered to be the collection, recording, organisation, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction of personal data.

PURPOSE

The term 'purpose' is a key concept in data protection regulation, defining the scope of the processing and assessing whether processing is lawful or not. The purpose refers to the aim pursued by the specific processing of personal data.

CONTROLLER

The controller is the natural or legal person who, alone or jointly with others, determines the purposes and means of the processing of personal data.

DATA SUBJECT

The data subject is generally defined as the person to whom the personal data relate.

PROCESSOR

The processor is the natural or legal person, public authority, agency or any other body that processes personal data on behalf of the controller. This will typically be a specialised third-party company entrusted by the controller to conduct the technical aspects of the processing, such as

the sorting or the combination of the personal data. The employee of the controller in charge of the security and management of the computer system is not to be considered as a processor.

THIRD PARTY

The third party is any natural or legal person, public authority, agency or any other body other than:

- the data subject,
- the controller,
- the processor
- and the persons who, under the direct authority of the controller or the processor, are authorised to process the data.

EUBIROD PIA Questionnaire

SECTION 1. Accountability for Personal Information

Accountability for Personal Information

Questions For Analysis		Yes	No	N/D or N/A	Provide Details
1.1	Has the custody and control of personal information been determined?				
1.2	Has the accountability of the registry/database custodian of personal information been documented?				
1.3	Are third parties involved in the custody or control of the personal information?				
1.4	If third parties are involved, do you have an agreement in place that establishes privacy requirements?				
1.5	Are there any requirements in registry/database legislation or policies on the management of personal information that affect the EUBIROD project?				

Discussion Points:

SECTION 2. Collection of Personal Information

Collection of Personal Information

Questions For Analysis	Yes	No	N/D or N/A	Provide Details
2.1 Do you have authority to collect personal information? Please indicate the authority. If there is no authority, please consult with your legal advisor to determine if there is authority to proceed.				
2.2 Is personal information being collected directly from the individual? If no, why not?				
2.3 Have the purposes for which the personal information is collected been documented? If yes, provide specifics				
2.4 Is all the personal information collected necessary to the registry/database?				
2.5 Are secondary uses contemplated for the information collected? If yes, describe them in the details column.				
2.6 If personal information is to be used or disclosed for a secondary purpose not previously identified, is consent required?				
2.7 If consent is not required for secondary purpose use or disclosure, is there authority for the use or disclosure (e.g. authorized by law)?				
2.8 Is information anonymized when used for planning, management and/or evaluation purposes?				
2.9 Is some personal information collected from a public database?				
2.10 Does the registry/database involve the collection from multiple sources through a common patient identifier? If yes, provide details about the identifier.				

Discussion Points:

SECTION 3. Consent

Consent

Questions For Analysis		Yes	No	N/D or N/A	Provide Details
3.1	Is consent obtained directly from the individual? If not, why not?				
3.2	Is consent required to collect and process data in the registry/database? If yes, how is consent obtained?				
3.3	Is consent clear and unambiguous?				
3.4	Can an individual refuse to consent to the collection or use of personal information for a secondary purpose, unless required by law?				
3.5	Are standards and mechanisms in place to ensure that the individual has capacity to give consent?				
3.6	Are standards and mechanisms in place to ensure the recognition of persons authorized to make decisions on behalf of others (e.g. a minor or incapacitated person)? If not why not?				

Discussion Points:

SECTION 4. Use of Personal Information

Use of Personal Information

Questions For Analysis	Yes	No	N/D or N/A	Provide Details
4.1 Do you have authority to use personal information? Please indicate the authority. If there is no authority please consult your legal advisor to determine the authority to proceed with the proposal.				
4.2 Is personal information used exclusively for the purpose for which the information was obtained or compiled?				
4.3 Are personal identifiers, such as a social insurance number, used for the purposes of linking across multiple databases?				
4.4 Where data matching, is it consistent with the stated purposes for which the personal information is collected?				
4.5 Does the data matching activity require a notification to the Privacy Commissioner?				

Discussion Points:

SECTION 5. Disclosure and Disposition of Personal Information

Disclosure and Disposition of Personal Information

Questions For Analysis	Yes	No	N/D or N/A	Provide Details
5.1 Is personal information disclosed with the consent of the individual?				
5.2 If personal information is not disclosed with consent, has the specific authority for disclosure been identified? If there is no authority to disclose personal information, please consult your departmental legal advisor.				
5.3 Are personal identifiers, such as a social insurance number, disclosed?				
5.4 Will personal information be processed, disclosed or retained outside the nation?				
5.5 Will personal information be disposed after a pre-determined time and, in all cases, when not anymore necessary for the purposes for which those data have been collected ?				

Discussion Points:

SECTION 6. Accuracy of Personal Information

Accuracy of Personal Information

Questions For Analysis	Yes	No	N/D or N/A	Provide Details
6.1 Will steps be taken to ensure that the personal information is accurate, complete and up-to-date?				
6.2 Does the record of personal information indicate the date of last information update?				
6.3 Is a record kept of the source of the information used to make changes?				
6.4 Where applicable, is there a procedure, automatically or at the request of an individual, to provide notices of correction to third parties to whom personal information has been previously disclosed?				
6.5 Is there a record kept with respect of requests for a review of errors or omissions & corrections or decisions not to correct?				
6.6 Is there a clearly defined process by which an individual may access, assess and discuss or dispute the accuracy of the record? Please briefly describe the steps?				

Discussion Points:

SECTION 7. Safeguarding Personal Information

Safeguarding Personal Information

Questions For Analysis	Yes	No	N/D or N/A	Provide Details
7.1 Have security procedures for the collection, transmission, storage and disposal of personal information, and access to it, been documented?				
7.2 Are program and information technology staff trained in the requirements for protecting personal information and are they aware of the relevant policies regarding breaches of security or confidentiality?				
7.3 Are there controls in place for any process to grant authorization to modify (add, change or delete) personal information from records?				
7.4 Are user accounts, access rights and security authorizations controlled by a system or record management process?				
7.5 Are security measures commensurate with the sensitivity of the information recorded?				
7.6 Are there contingency plans and documented procedures in place to identify and respond to security breaches or disclosures of personal information in error?				
7.7 Are there documented procedures in place to communicate security violations to the data subject, law enforcement authorities and relevant program managers?				
7.8 Is there a plan for quality assurance and audit programs to assess the ongoing state of the safeguards applicable to the system?				

Discussion Points:

SECTION 8. Openness

Openness

Questions For Analysis	Yes	No	N/D or N/A	Provide Details
8.1 Is there a communication plan to explain to the public how personal information will be managed and protected?				
8.2 Is there a clearly defined and easy process for individuals to access such information and/or communicate with appropriate individuals with respect to policies and practices relating to management and protection of personal information?				

Discussion Points:

SECTION 9. Individual's Access to Personal Information

Individual's Access to Personal Information

Questions For Analysis	Yes	No	N/D or N/A	Provide Details
9.1 Is the system designed to ensure that an individual can have access to his/her personal information?				
9.2 Is the system designed to ensure that an individual has been notified that a correction to his/her information has been made?				
9.3 Are all custodians and participants aware of an individual's right of access and the complaint process?				
9.4 Has consideration been given to providing individuals "routine" access to their personal information?				

Discussion Points:

SECTION 10. Challenging Compliance

Challenging Compliance

Questions For Analysis	Yes	No	N/D or N/A	Provide Details
10.1 Are the complaint procedures implemented in the registry/database consistent with legislated requirements?				
10.2 Are there oversight and review mechanisms implemented or available to ensure accountability?				
10.3 Have oversight agencies, including the Office of the Privacy Commissioner, issued reports or opinions on issues that would be relevant to the project? If yes, please provide a summary of the above in the details column and append to final report.				

Discussion Points:

SECTION 11. Anonymisation Process for Secondary Uses of Health Data

Anonymisation Process

Questions For Analysis	Yes	No	N/D or N/A	Provide Details
1. Is a standard anonymisation procedure envisaged by your Centre/local/national regulation or rule before further processing data for secondary uses?				
2. If yes, is the applied procedure compliant with international technical standards and continuously updated according to the state of the art?				
3. If yes, is the anonymisation process performed in compliance with the Data Protection Principles; for instance, performed confidentially, providing information to patients about the processing operation, applying security mechanisms for data storage and retention, etc.?				

Please provide a detailed description of the anonymisation process, including materials, methods and techniques used.

Appendix 2. Statistical Source Code

```
rm(list=ls()) # clean environment

library(lattice) # load lattice library
library(reshape) # load library to make recoding easy
library(prettyR) # load tables library
library(Cairo)   # load PDF routines
library(nlme)    # load non linear multivariate models

launchtime<-format(Sys.time(),"%d%m%y%H%M%S")

# MAIN FUNCTION

biro_pia<-function(homedir,infile,outdir) {

# create directory structure
dir.create(paste(homedir,outdir,sep=""),
           showWarnings = FALSE, recursive = TRUE)
dir.create(paste(homedir,outdir,"/questions/",sep=""),
           showWarnings = FALSE, recursive = TRUE)
dir.create(paste(homedir,outdir,"/recoded_questions/",sep=""),
           showWarnings=FALSE,recursive = TRUE)
dir.create(paste(homedir,outdir,"/absolutefactors/",sep=""),
           showWarnings=FALSE,recursive = TRUE)
dir.create(paste(homedir,outdir,"/scaledfactors/",sep=""),
           showWarnings=FALSE, recursive = TRUE)
dir.create(paste(homedir,outdir,"/echarts/",sep=""),showWarnings=FALSE,recursive= TRUE)

#read master index dataset and loads into memory
pia<-read.table(paste(homedir,infile,sep=""),header=TRUE,sep=",",na.strings="")

# Print Histograms of all original variables

for (j in 2:dim(pia)[2]) {
 file <- paste(homedir,outdir,"/questions/",names(pia)[j],".png",sep="")
 png(file)
 print( histogram(as.factor(pia[,j]),
        type=c("count"),
        labels=FALSE,
        freq=TRUE,
        col="lightblue",
        main = paste("Histogram of",names(pia[j])),
        xlab="Responses",border="blue")
       )
dev.off()

}

# Recode based on scoring system

# for all variables missing equals ND/NA
for (j in 2:dim(pia)[2]) {
 pia[,j][pia[,j]==9]<-0
}

# Section 1
```

```
pia$X1_1[pia$X1_1==2]<-0
pia$X1_2[pia$X1_2==2]<-0
pia$X1_34<-1
pia$X1_34[pia$X1_3==1 & (pia$X1_4!=1)]<-0

# Section 2
pia$X2_1[pia$X2_1==2]<-0
pia$X2_2[pia$X2_2==2]<-0
pia$X2_3[pia$X2_3==2]<-0
pia$X2_4[pia$X2_4==2]<-0
pia$X2_567<-1
pia$X2_567[(pia$X2_6==0 | pia$X2_7==0)]<-0
pia$X2_8[pia$X2_8==2]<-0

# Section 3
pia$X3_1[pia$X3_1==2]<-1
pia$X3_2[pia$X3_2==0]<-1
pia$X3_2[pia$X3_2==2]<-0
pia$X3_3[pia$X3_3==2]<-0
pia$X3_4[pia$X3_4==2]<-0
pia$X3_5[pia$X3_5==2]<-0
pia$X3_6[pia$X3_6==2]<-0

# Section 4
pia$X4_1[pia$X4_1==2]<-0
pia$X4_2[pia$X4_2==2]<-0
pia$X4_3[pia$X4_3==0]<-1
pia$X4_3[pia$X4_3==1]<-0
pia$X4_3[pia$X4_3==2]<-1
pia$X4_4[pia$X4_4==2]<-1

# Section 5
pia$X5_1[pia$X5_1==2]<-1
pia$X5_2[pia$X5_2==2]<-1
pia$X5_3[pia$X5_3==0]<-1
pia$X5_3[pia$X5_3==1]<-0
pia$X5_3[pia$X5_3==2]<-1
pia$X5_4[pia$X5_4==0]<-1
pia$X5_4[pia$X5_4==1]<-0
pia$X5_4[pia$X5_4==2]<-0
pia$X5_5[pia$X5_5==2]<-0

# Section 6
pia$X6_1[pia$X6_1==2]<-0
pia$X6_2[pia$X6_2==2]<-1
pia$X6_3[pia$X6_3==2]<-1
pia$X6_4[pia$X6_4==2]<-1
pia$X6_5[pia$X6_5==2]<-1
pia$X6_6[pia$X6_6==2]<-0

# Section 7
pia$X7_1[pia$X7_1==2]<-0
pia$X7_2[pia$X7_2==2]<-0
pia$X7_3[pia$X7_3==2]<-0
pia$X7_4[pia$X7_4==2]<-0
pia$X7_5[pia$X7_5==2]<-0
pia$X7_6[pia$X7_6==2]<-0
pia$X7_7[pia$X7_7==2]<-0
pia$X7_8[pia$X7_8==2]<-0

# Section 8
```

```
pia$X8_1[pia$X8_1==2]<-1
pia$X8_2[pia$X8_2==2]<-1

# Section 9
pia$X9_1[pia$X9_1==2]<-1
pia$X9_2[pia$X9_2==2]<-1
pia$X9_3[pia$X9_3==2]<-1
pia$X9_4[pia$X9_4==2]<-1

# Section 10
pia$X10_1[pia$X10_1==2]<-1
pia$X10_2[pia$X10_2==2]<-1

# Section 11
pia$X11_1[pia$X11_1==2]<-1
pia$X11_2[pia$X11_2==2]<-1
pia$X11_3[pia$X11_3==2]<-1

# Factors

# Initialize N=11 factors

factors<-pia[1]

factors$A1=pia$X1_1+pia$X1_2+pia$X1_34
factors$A2=pia$X2_1+pia$X2_2+pia$X2_3+pia$X2_4+pia$X2_567+pia$X2_8
factors$A3=pia$X3_1+pia$X3_2+pia$X3_3+pia$X3_4+pia$X3_5+pia$X3_6
factors$A4=pia$X4_1+pia$X4_2+pia$X4_3+pia$X4_4
factors$A5=pia$X5_1+pia$X5_2+pia$X5_3+pia$X5_4+pia$X5_5
factors$A6=pia$X6_1+pia$X6_2+pia$X6_3+pia$X6_4+pia$X6_5+pia$X6_6
factors$A7=pia$X7_1+pia$X7_2+pia$X7_3+pia$X7_4+pia$X7_5+pia$X7_6+pia$X7_7+pia$X7_8
factors$A8=pia$X8_1+pia$X8_2
factors$A9=pia$X9_1+pia$X9_2+pia$X9_3+pia$X9_4
factors$A10=pia$X10_1+pia$X10_2
factors$A11=pia$X11_1+pia$X11_2+pia$X11_3

factors$total<-
factors$A1+factors$A2+factors$A3+factors$A4+factors$A5+factors$A6+factors$A7+factors$A8+
factors$A9+factors$A10+factors$A11

labels=c("Accountability","Collection","Consent","Use","Disclosure","Accuracy","Safeguar
ding","Openness","Access","Compliance","Anonymisation","TOTAL")
labs=c("A1","A2","A3","A4","A5","A6","A7","A8","A9","A10","A11","TOTAL")

# Compute overall score

score<-pia[1]
score$A1<-factors$A1/3*100
score$A2<-factors$A2/6*100
score$A3<-factors$A3/6*100
score$A4<-factors$A4/4*100
score$A5<-factors$A5/5*100
score$A6<-factors$A6/6*100
score$A7<-factors$A7/8*100
score$A8<-factors$A8/2*100
score$A9<-factors$A9/4*100
score$A10<-factors$A10/2*100
score$A11<-factors$A11/3*100
```

```
score$total<-
(score$A1+score$A2+score$A3+score$A4+score$A5+score$A6+score$A7+score$A8+score$A9+score$
A10+score$A11)/11

scorescaled<-score

scoreband<-pia[1]
scoreband$A1<-"D"
scoreband$A2<-"D"
scoreband$A3<-"D"
scoreband$A4<-"D"
scoreband$A5<-"D"
scoreband$A6<-"D"
scoreband$A7<-"D"
scoreband$A8<-"D"
scoreband$A9<-"D"
scoreband$A10<-"D"
scoreband$A11<-"D"
scoreband$total<-"D"

for (j in 2:13) {

scorescaled[j]=score[j]/100

scoreband[j][score[j]>40 & score[j]<=60]<-"C"
scoreband[j][score[j]>60 & score[j]<=80]<-"B"
scoreband[j][score[j]>80 & score[j]<=100]<-"A"

}

# Save recoded values in csv format
write.csv(pia,file=paste(homedir,outdir,"/recoded_questions.csv",sep=""))
write.csv(factors,file=paste(homedir,outdir,"/factors.csv",sep=""))
write.csv(score,file=paste(homedir,outdir,"/scores.csv",sep=""))
write.csv(scoreband,file=paste(homedir,outdir,"/scorebands.csv",sep=""))

# Histograms of all original variables

for (j in 2:dim(pia)[2]) {
 file <- paste(homedir,outdir,"/recoded_questions/",names(pia)[j],".png",sep="")
 png(file)
 print( histogram(as.factor(pia[,j]),
        type=c("count"),
        labels=FALSE,
        freq=TRUE,
        col="lightblue",
        main = paste("Histogram of",names(pia[j])),
        xlab="Responses",border="blue")
       )
dev.off()

}

# Histograms of individual factors

i<-1
for (j in 2:12) {
 file <- paste(homedir,outdir,"/absolutefactors/",labs[i],".png",sep="")
 png(file)
 print( histogram(as.factor(factors[,j]),
```

```
        type=c("count"),
        labels=FALSE,
        freq=TRUE,
        col="lightblue",
        main = paste("Histogram of" ,labels[i]),
        xlab="Responses",border="blue")
       )
 dev.off()
 i<-i+1
}

# Histograms of Total Score
file <- paste(homedir,outdir,"/scaledfactors/total.png",sep="")
png(file)
print( histogram(score$total,
        xlim=c(-10,110),
        type=c("count"),
        labels=FALSE,
        freq=TRUE,
        col="lightblue",
        main ="Histogram of Total Score",
        xlab="Responses",
        border="blue")
      )

dev.off()

# Boxplots
file <- paste(homedir,outdir,"/scaledfactors/A_box.png",sep="")
png(file)
print( boxplot(score[2:12],
       type=c("count"),
       labels=FALSE,
       freq=TRUE,
       col="lightblue",
       main = paste("Boxplot of all Factors"),
       xlab="Responses",border="blue")
      )

dev.off()

# Reorder outputs for Starplots
scorescaledtwo<-scorescaled[c(1,4,3,2,12,11,10,9,8,7,6,5)]

file <- paste(homedir,outdir,"/scaledfactors/A_star.png",sep="")
png(file)
print(stars(scorescaledtwo[2:12],
      key.xpd=-1,
      key.loc=c(9.2,2.25),
      xpd=TRUE,
      lwd=1.5,
      len=0.8,
      labels=pia$COUNTRY,
      main="PIA Factors by Diabetes Register",
      scale=FALSE)
     )

dev.off()

# heatmap
file <- paste(homedir,outdir,"/scaledfactors/A_heatmap.png",sep="")
```

```
png(file)
print( heatmap(as.matrix(scorescaled[2:12])) )

dev.off()

# Privacy Self Evaluation Chart

index<-0;factor<-"A";value<-0;lcl<-0;average<-0;ucl<-0
new.guy<-data.frame(index,factor,value,lcl,average,ucl)

for (k in 1:dim(pia)[1]) {

score_ci<-
data.frame(index=numeric(0),factor=character(0),value=numeric(0),lcl=numeric(0),average=
numeric(0),ucl=numeric(0))

 for (j in 2:13) {

  new.guy$index<-(j-1)
  new.guy$factor<-labels[j-1]
  new.guy$value<-score[k,j]
  new.guy$lcl<-mean(score[j])-1.96*sd(score[j])/sqrt(18)
  new.guy$average<-mean(score[j])
  new.guy$ucl<-mean(score[j])+1.96*sd(score[j])/sqrt(18)

  score_ci<-rbind(score_ci,new.guy)

 }

# Statistical Tables

 print(score_ci)
 print(describe(scorescaled))
 print(summary(scorescaled))

 for (j in 3:6) {

  score_ci[j][score_ci[j]<0]<-0
  score_ci[j][score_ci[j]>100]<-100

 }

 file <- paste(homedir,outdir,"/echarts/echart_",factors[k,1],".png",sep="")
 png(file)

 print(dotplot(~ value + lcl + average + ucl | factor,
         data=score_ci,
         xlim=c(-10,110),
         main=paste("Register:",factors[k,1],"- Privacy Self Evaluation Chart"),
         xlab="Scaled Score",
         ylab=NULL,
         index.cond=list(c(11,4,6,1,9,10,3,8,12,7,5,2)),
         auto.key=list(columns=4),
         par.settings=list(superpose.symbol=list(col=c(1,2,1,2),
         pch=c(23,5,3,5),
         cex=c(0.75,1.3,0.4,1.3))),
         layout=c(1,12)))

 dev.off()
```

```
}

}

# Launch software on EUBIROD PIA Questionnaire data

biro_pia(homedir="/home/pia_d5.2_250610/pia_statistics_310510",
         infile="/eubirod_pia_questionnaire.csv",
         outdir=paste("/pia_graphs_#",launchtime,"/images",sep=""))
```

References

1 DG SANCO Task Force of Major and Chronic Diseases, Major and Chronic diseases in the European Union - Report 2007, European Commission, Luxembourg, 2008; available at: http://ec.europa.eu/health/ph_threats/non_com/docs/mcd_report_en.pdf
2 Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, Official Journal of the European Communities No. L 281/31; available at: http://ec.europa.eu/justice_home/fsj/privacy/law/index_en.htm
3 Treaty of Lisbon amending the Treaty on European Union and the Treaty establishing the European Community, signed at Lisbon, 13 December 2007. Official Journal of the European Union, 2007/C 306/01; available at: http://eur-lex.europa.eu/JOHtml.do?uri=OJ:C:2007:306:SOM:EN:HTML
4 Davies S. Michael, Privacy and human rights [book review]. Web Journal of Current Legal Issues. Blackstone Press, 1995; available at: http://webjcli.ncl.ac.uk/articles1/davies1.html
5 Warren S, Brandeis L. The Right to Privacy. Harvard Law Review 1890; 4:193–220
6 Calcutt D QC .(Chairman). Report of the Committee on Privacy and Related Matters. London: Cmnd. 11027, 1990
7 Universal Declaration of Human Rights, adopted and proclaimed by General Assembly resolution 217 A (III) of December 10, 1948; available at: http://www.un.org/Overview/rights.html
8 International Covenant on Civil and Political Rights, adopted and opened for signature, ratification and accession by General Assembly resolution 2200A (XXI) of December 16, 1966, entry into force March 23rd 1976; available at: http://www.unhchr.ch/html/menu3/b/a_ccpr.htm
9 International Convention on the Protection of the Rights of All Migrant Workers and Members of Their Families, adopted by General Assembly resolution 45/158 of December 18, 1990; available at: http://www.unhchr.ch/html/menu3/b/m_mwctoc.htm
10 Convention on the Rights of the Child, adopted and opened for signature, ratification and accession by General Assembly resolution 44/25 of November 20, 1989, entry into force September 2, 1990; available at: http://www.unhchr.ch/html/menu3/b/k2crc.htm
11 Council of Europe, Convention for the Protection of Human Rights and Fundamental Freedoms, (ETS no: 005) open for signature November 4, 1950, entry into force September 3, 1950; available at: http://conventions.coe.int/Treaty/en/Treaties/Html/005.htm
12 Strossen N, Recent United States and Intl. Judicial Protection of Individual Rights: A comparative Legal Process Analysis and Proposed Synthesis. Hastings Law Journal 1990; 41: 805
13 European Court of Human Rights, Case of Klass and Others: Judgement of 6 September 1978, Series A No. 28 (1979). Malone v. Commissioner of Police, Series A82 (1984); available at: http://hudoc.echr.coe.int/hudoc/ViewRoot.asp?Item=5&Action=Html&X=1007095902&Notice=0&Noticemode=&RelatedMode=0;
14 European Court of Human Rights, Leander v. Sweden, series A No 116 (1987); available at: http://hudoc.echr.coe.int/hudoc/ViewRoot.asp?Item=0&Action=Html&X=1007101431&Notice=0&Noticemode=&RelatedMode=0
15 Council of Europe. Convention for the protection of individuals with regard to automatic processing of personal data. Strasbourg: The Council, 1981; available at: http://conventions.coe.int/Treaty/en/Treaties/Html/108.htm

16 OECD, Guidelines Governing the Protection of Privacy and Transborder Data Flows of Personal Data. Paris, 1981; available at: http://www.oecd.org/dsti/sti/it/secur/prod/PRIV-EN.HTM
17 Council of Europe Convention on Human rights and Biomedicine (Oviedo 1997); available at: http://conventions.coe.int/Treaty/EN/Treaties/Html/164.htm
18 Charter of Fundamental Rights of the European Union (2000/C 364/01); available at: http://ec.europa.eu/justice_home/unit/charte/index_en.html
19 Official Journal of the European Union C 310 Volume 47 of 16 December 2004; available at: http://eur-lex.europa.eu/JOHtml.do?uri=OJ:C:2004:310:SOM:EN:HTML
20 Treaty of Lisbon amending the Treaty on European Union and the Treaty establishing the European Community, signed at Lisbon, 13 December 2007. Official Journal of the European Union. 2007/C 306/01; available at: http://eur-lex.europa.eu/JOHtml.do?uri=OJ:C:2007:306:SOM:EN:HTML
21 Additional Protocol to the Convention on Human Rights and Biomedicine, concerning Biomedical Research. Strasbourg, 25.I.2005; available at: http://conventions.coe.int/Treaty/EN/Treaties/Html/195.htm
22 McClelland R at all. European Standards on Confidentiality and Privacy in Healthcare, EuroSOCAP Project (2003-2006) available at: http://www.orpha.net/testor/doc/july05/EuroSOCAP.pdf
23 McClelland R at all. European Standards on Confidentiality and Privacy in Healthcare, EuroSOCAP Project (2003-2006); available at: http://www.orpha.net/testor/doc/july05/EuroSOCAP.pdf
24 European Commission Contract 30-CE-0041734/00-55, European Health Management Association, Legally eHealth, 2006, D.2(PUB) v. 8; available at: http://ec.europa.eu/information_society/activities/health/docs/studies/legally-ehealth-report.pdf
25 Council of Europe Committee of Ministers, Recommendation No R (97) 5 of the Committee of Ministers to Member States on the Protection of Medical Data; available at: https://wcd.coe.int/com.instranet.InstraServlet?command=com.instranet.CmdBlobGet&InstranetImage=564487&SecMode=1&DocId=560582&Usage=2
26 EUropean Best Information through Regional Outcomes in Diabetes (EUBIROD), available at:
http://www.eubirod.eu/
27 BIRO Consortium (2009), Best information through regional outcomes: a shared European diabetes information system for policy and practice, Università di Perugia, Perugia, Italia, available at: http://www.eubirod.eu/documents/downloads/BIRO_Monograph.pdf.
28 Di Iorio CT, Carinci F, Azzopardi J, Baglioni V, Beck P, Cunningham S, Evripidou A, Leese G, Loevaas KF, Olympios G, Orsini Federici M, Pruna S, Palladino P, Skeie S, Taverner P, Traynor V, Massi Benedetti M (2009) Privacy impact assessment in the design of transnational public health information systems: the BIRO project, Journal of Medical Ethics, Dec;35(12):753-61.
29 Concil of Europe, Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by Protocol No. 11, Rome (1950), available at: http://www.echr.coe.int/nr/rdonlyres/d5cc24a7-dc13-4318-b457-5c9014916d7a/0/englishanglais.pdf
30 Beck P, Truskaller T, Rakovac I, Bruner F, Zanettin D, Pieber TR., Information systems for administration, clinical documentation and quality assurance in an Austrian disease management programme, Stud Health Technol Inform. 2009;150:379-83.
31 Debacker N, Nobels F, Vandenberghe H, Van Crombrugge P, Scheen A, Van Casteren V., Organization of a quality-assurance project in all Belgian multidisciplinary diabetes centres

treating insulin-treated diabetes patients: 5 years' experience, Diabet Med. 2008 Feb;25(2):179-85.
32 Poljicanin T, Pavlić-Renar I, Metelko Z., CroDiab NET, electronic diabetes registry, Acta Med Croatica. 2005;59(3):185-9.
33 Széles G, Vokó Z, Jenei T, Kardos L, Pocsai Z, Bajtay A, Papp E, Pásti G, Kósa Z, Molnár I, Lun K, Adány R., A preliminary evaluation of a health monitoring programme in Hungary, Eur J Public Health. 2005 Feb;15(1):26-32.
34 Benedetti MM, Carinci F, Federici MO, The Umbria diabetes register, Diabetes Res Clin Pract. 2006 Dec;74 Suppl 2:S200-4.
35 de Beaufort CE, Reunanen A, Raleigh V, Storms F, Kleinebreil L, Gallego R, Giorda C, Midthjell K, Jecht M, de Leeuw I, Schober E, Boran G, Tolis G., European Union diabetes indicators: fact or fiction? Eur J Public Health. 2003 Sep;13(3 Suppl):51-4.
36 Nijpels G, Hoovers T, Dekker JM, Heine RJ. Regionaal georganiseerde diabeteszorg. Het Diabetes Zorgsysteem West-Friesland. Medisch Contact. 1998;53:1164–1166.
37 Jarosz-Chobot P, Deja G, Polanska J. Epidemiology of type 1 diabetes among Silesian children aged 0-14 years, 1989-2005, Acta Diabetol. 2010 Mar;47(1):29-33. Epub 2009 Jan 31.
38 Pruna S, Stanciu E, Macarie A, Pruna A, Ionescu-Tirgoviste C. A networked electronic patient record system for diabetes, Stud Health Technol Inform. 2002;90:282-7.
39 Morris AD, Boyle DI, MacAlpine R, Emslie-Smith A, Jung RT, Newton RW, MacDonald TM., The diabetes audit and research in Tayside Scotland (DARTS) study: electronic record linkage to create a diabetes register. DARTS/MEMO Collaboration, BMJ. 1997 Aug 30;315(7107):524-8.
40 Hjerpe P, Merlo J, Ohlsson H, Bengtsson Boström K, Lindblad U. Validity of registration of ICD codes and prescriptions in a research database in Swedish primary care: a cross-sectional study in Skaraborg primary care database, BMC Med Inform Decis Mak. 2010 Apr 23;10:23
41 Bratina NU, Tahirović H, Battelino T, Krzisnik C., Incidence of childhood-onset Type I diabetes in Slovenia and the Tuzia region (Bosnia and Herzegovina) in the period 1990-1998. Diabetologia. 2001 Oct;44 Suppl 3:B27-31.
42 Treasury Board of Canada Secretariat, Privacy Impact Assessment Guidelines: A Framework to Manage Privacy Risks, Questionnaire A: For Federal Programs and Services, 2002. Available at: http://www.tbs-sct.gc.ca/pubs_pol/ciopubs/pia-pefr/paipg-pefrld01-eng.asp#5.
43 R Development Core Team (2010), R: A Language and Environment for Statistical Computing, available at: http://cran.r-project.org/doc/manuals/refman.pdf
44 Roos LL, Menec V, Currie RJ. Policy analysis in an information-rich environment, Soc Sci Med. 2004 Jun;58(11):2231-41
45 Roos LL, Gupta S, Soodeen RA, Jebamani L. Data quality in an information-rich environment: Canada as an example, Can J Aging. 2005 Spring;24 Suppl 1:153-70
46 Holman CD, Bass AJ, Rosman DL, Smith MB, Semmens JB, Glasson EJ, Brook EL, Trutwein B, Rouse IL, Watson CR, de Klerk NH, Stanley FJ, A decade of data linkage in Western Australia: strategic design, applications and benefits of the WA data linkage system, Aust Health Rev. 2008 Nov;32(4):766-77.
47 Ingelfinger J, Drazen J. Registry research and Medical Privacy. N Engl J Med, 2004;350:1452–53
48 Marieke Verschuuren, Gérard Badeyan et Al, The European data protection legislation and its consequences for public health monitoring: a plea for action, Eur J Public Health, 2008 December; 18(6): 550–551; available at: http://eurpub.oxfordjournals.org/cgi/reprint/18/6/550.pdf
49 Mati Rahu, Martin McKee, Epidemiological research labelled as a violation of privacy: the case of Estonia, International Journal of Epidemiology 2008;37:678–682; available at:

http://ije.oxfordjournals.org/cgi/reprint/37/3/678