Policy versus privacy: an analysis of ethical issues in a health monitoring project

by Lies Lammens & Patrick Deboosere, Vrije Universiteit Brussel, Belgium

Abstract

The objective of this paper is to increase awareness about ethical issues in demographic health research, recently being prominent due to new technological research possibilities. More specifically, we examine the tension between the demand for knowledge in our societies, and the protection of people's privacy.

On the basis of the example of a health monitoring project currently under construction in Belgium, we propose a model which could be used as a framework for the ethical evaluation of data collection projects with the purpose of health monitoring. Crucial in our model is the dimension of time. Ethical implications on a short as well as on a long term are being considered.

Even though there is no ideal solution to overcome the ethical dilemma discussed here, we believe that an extensive health data collection can be compatible with the protection of people's privacy on the short and long term. We argue that democratic control is crucial for protection against a potential abuse of personal data on a large scale.

1. Research in modern societies

1.1. Technological and societal context

Modern societies are faced with increasingly complex, often interrelated, policy issues in areas such as social security, environment, health care, population aging, changing family structures and poverty reduction. Nowadays specific questions arise such as how changes in the age structure of the population affect the health care system, and how this interacts with various social security domains. Addressing these kind of issues requires the collection of detailed data on the state of the population in various domains, and sophisticated research methods to analyze and interpret these data.

Already in premodern times, the population size was calculated and used for taxation purposes, military service or to estimate economic capacity. But only after the development of modern administration, technology, and professional expertise, large-scale population databases could be set up, thus allowing a systematic collection of uniform, periodic information about a nation's population, (Seltzer & Anderson, 2001, p.481).

Technological developments have dramatically altered the traditional research environment. For example, more and more electronic means are being used to acquire, use and store information, compared to paper records used in the past. Electronic information offers many advantages: it can be standardized more easily. Moreover it

increases data reliability and access, reduces errors, saves money, and improves the data collection speed. In principle, electronic data can also be secured better than paper records, since they facilitate authentification, authorization, auditing, and accountability processes (Myers et al, 2008, pp.793-794). Another example of an important advance in information technology is the development of the Internet. Its increasingly widespread use in daily life has resulted in information being transferred globally and more rapidly than ever before. Besides that, more sophisticated statistical techniques, data analyzing programs and search strategies have been developed, and have been used to manage the information acquired.

Consequently, not only have the costs of data collection, data storage, data analysis, data integration and data dispersion decreased in recent years. Also, the new IT possibilities have enabled a more systematic and detailed monitoring of the population, and have thus lead to a better knowledge of the population state. A better follow-up of relevant population indicators can guide and validate policies for the common good. In fact, data are the necessary base for the direction and implementation of accurate policies (Duncan, 2004, p.4).

Nevertheless, these technology advances impose a new, great challenge to our societies: people's privacy rights might get endangered. Privacy concerns have been around at least for 50 years, i.e. ever since the arrival of large electronic databases. Despite the many advantages, electronic information can also be more easily copied, transported and spread in comparison to paper records. Thus, if misuse happens, the consequences may be much bigger with electronic data than with data on paper (Myers et al, 2008, p.794).

Even though data might be protected more carefully nowadays, mainly through more advanced technological security measures as well as well-defined legal frameworks, it seems like no system can really prevent misuse to happen. It suffices to consider recent events of data losses or data breaches, to realize that complete exclusion of data misuse is very difficult, if not impossible. The internationally increasing amount of hackings of 'secured' credit cards (De Morgen, 2008); the accidental attachment of an electronic file containing the names and addresses of 6500 HIV/AIDS patients to an e-mail (Myers et al, 2008, p.793); the loss of secret details on the war against drugs in Colombia by a British spy (De Morgen, 2009); a theft from an employee's car of a state health department laptop computer containing information on approximately 1600 families (ConsumerUnion, 2005) are just some examples of recent data thefts or losses, showing the vulnerability of our current 'highly secured' system. It seems like more sophisticated research possibilities inevitably carry the risk of personal information being disclosed, despite efforts to keep it confidential.

Not only the technological context, but also the societal context of data collection and exploitation in modern societies, has recently changed. Since 9/11 governmental control on potential 'national security threats' has become more important against the protection of people's privacy. Thus the September 11 attacks remarkably changed the role of information and the way people perceive privacy rights in our society (Duncan, 2004, pp.4-5). We cite for example the introduction, after 9/11, of biometric passports in the US

and subsequently in other countries. These new passports contain a chip carrying digital images of the bearer's face and fingerprints. Privacy concerns contain the idea that the personal information on these passports will be collected in databases and used in unforeseen ways in the future (Rue 89, 2008).

In different countries so-called 'super databases' are developing, like Edvige in France or ANG in Belgium. They collect an enormous amount of, even very personal, data on people's lives, behaviors and networks and are meant to prevent the public order to be broken. The police database Edvige for example was about to be created in France to store data on anyone aged 13 or above, who is "likely to breach public order". It includes 'common' data like individuals' occupation, address, phone number, pictures, car number plate, fiscal data, and judicial antecedents, but also 'sensitive' data like racial or ethnic origin, political, philosophical or religious opinions, information on social networks, membership of political parties or trade unions, health status and sexual activities (International Herald Tribune, 2008). Protest against Edvige from the French public, expressed by a petition signed by a great amount of individuals, forced the government to withdraw the decree regulating the framework of the database.

1.2. Health research

One of the research domains facing the dilemma between an extensive population monitoring and an adequate protection of individuals' privacy is the study of population health, somewhere at the crossroads between epidemiology and health demography.

Healthcare is given a great amount of attention by governments in modern societies. The healthcare sector nowadays receives the largest share of government expenditures, and that share continues to grow. Between 1995 and 2006, average OECD health expenditure per capita has grown annually by around 4%. Average economic growth over the same period was 2.5%. In 2006 the OECD average of total health expenditures was 10.5% of the Net National Income (NNI). Behind this OECD average, significant variations can be observed both between countries and over time. In Belgium, the total amount of healthcare expenses has risen to almost € 32 billion in 2006, which counted for approximately 12% of the NNI. Compared to 1995, this represents an increase of 4%. Of all OECD countries, the United States spends the largest share of its NNI on health expenses, i.e. approximately 17% in 2006, which represents a growth of 3.4% compared to 1995. Countries with the lowest proportion of health expenditures, include Turkey (around 4.5% in 2005), and Korea, Poland and Mexico (around 7.5% in 2006) (OECD, 2009, pp.116-117). For state and local public health departments, assessing population health has clearly become a major concern. The more so as expenses will inevitably keep on rising the next decennia.

On-going debate concerning rising healthcare costs, recognizes the importance of the impact of the ageing of the population. Some research results point to other factors such as new medical technologies and services, broader-access health plans, increased consumer demand, increased use and cost of pharmaceuticals, increasingly unhealthy

lifestyles and the (private/public) organization of the healthcare system. To be able to accurately assess the impact of the ageing population and other factors responsible for rising health care costs, as well as to react with the most adequate policy strategies, societies are forced to systematically follow up the health state of their population. That requires the acquisition, use and storage of detailed health related information on individuals.

As a consequence, the study of population health has strongly intensified in recent decades. Moreover, the complexity of current health policy issues together with the new technological context has brought along a particular way of collecting and storing personal health information.

Firstly, research is no longer conducted only on the basis of databases containing data on a certain proportion of the population, such as people being questioned for a particular survey. Rather, the focus is on databases covering the whole population. The records may be based on citizen-government interactions – when people are for example obliged to report certain information to the government to be entitled to certain services -, on government-mandated data provision from organisations, or - as in censuses - on data obtained directly from respondents.

Most of the data provision for this kind of databases is mandated by legislation and is obtained through systems of administrative records. There are many advantages linked to the use of administrative data. For example, they give researchers access to information that individuals may not be able to recall or estimate accurately in a survey context. Survey data can be biased as a result of flaws in respondents' memory or their understanding of measurement concepts (Mackie & Bradburn, 2000, p.7). Besides, administrative data are, if certain data conditions are fulfilled, cost-efficient. On the other hand though, administrative data might entail very personal data to be collected and used for research purposes without people knowing or having been informed about their data being used.

Secondly, more records get linked across databases, eventually covering different domains, on the basis of unique identifiers. Linking makes it possible to get more out of isolated datasets that would otherwise have limited application. Health survey data can for example be linked to information from health records - collected from e.g. people in hospitals, emergency rooms and doctors' offices -, which facilitates a broad spectrum of research that could otherwise not be conducted. Through linkage of these two sources, it can for example be examined which lifestyles are associated with which health conditions, or which treatments have worked with which particular background or lifestyle factors (Fellegi, 2004, p.143).

The linking process can increase the value of datasets, reduce data collection redundancies, and improve data accuracy in a cost-effective manner. Additionally, linking facilitates research on infrequent events, such as rare diseases, that affect only a small percentage of the population, since in such cases, working from general sample

data does not provide adequate sample sizes for target groups (Mackie & Bradburn, 2000, pp.6-8).

But a challenge on linking individual data, requiring access to micro data, is the increased risk of identification of the individuals concerned. The linking process poses a great challenge in obtaining 'informed consent', i.e. the right to give or deny consent for the use of information about oneself, the 'classical' solution to overcome the 'knowledge versus privacy'-dilemma. While informed consent is relatively easy to handle in (voluntary) surveys, new research methodologies - including the use of administrative data and record linkage - are ultimately challenging this form of protection. Besides, it is impossible to foresee the statistical analyses through linkage that might become desirable to be carried out in the future (Fellegi, 2004, p.146).

Thirdly, there is an increasing interest in analyzing individual life histories by exploiting longitudinal data, which require measuring characteristics of the same study objects at least at two points in time. The collection of longitudinal data enables us to follow individuals over time, and to monitor changes during their life course. Longitudinal research differs from the collection of cross-sectional data, in which the observed information is representative of the population at a certain moment in time and the temporal aspect of a specific individual's journey is usually not available. Longitudinal data have considerable advantages over more widely available cross-sectional data for social science analysis. They permit among others tracing the dynamics of certain behaviour and identifying the influence of past behaviours on current behaviours (Alderman et al, 2001, p.83). Longitudinal research on poverty dynamics succeeded for example in replacing longstanding beliefs about the permanence of poverty with knowledge about the extent to which poverty is both widespread and temporary for a large proportion of the population (Duncan & Pearson, 1991, p.220). Besides, in longitudinal analyses, by identifying observations on the same individuals over a period of time, it is possible to focus on changes occurring within subjects and make population inferences that are not as sensitive to between-subject variation (Yee & Niemeier, 1996, p.1). These advantages are substantial when studying processes occurring over time and trying to relate social outcomes to underlying causes.

Despite the various benefits offered by longitudinal research, the collection of longitudinal data is likely to be difficult and expensive. Besides that, respecting statistical confidentiality while analysing data from this perspective, is a particularly difficult challenge. Longitudinal data contain more information on the characteristics and behavior of individuals than cross-sectional data. Even if anonymized, the risk of disclosure is higher. Moreover, longitudinal data have traditionally been collected by means of surveys, questioning the same sample of respondents at different points in time. Recently though, longitudinal data can be acquired 'more easily' through (repeated) linkage of administrative individual data. This increases the risk of identification considerably.

1.3. Illustration: the Belgian eHealth-platform

A concrete initiative undertaken by many governments being confronted with an increasingly complex health care system is the development of a national infrastructure of health information. Characteristic features of such an infrastructure are the existence of electronic patient records, databases enabling more comprehensive and systematic collection, electronic patient cards enabling patient data to be recorded on, unique personal identifiers, internal networks designed to share information among affiliated organizations that provide medical services, reimbursement services, and pharmaceutical agents, and public on-line networks that allow clinicians, researchers, and health care managers to share information (Gostin, 1997, p.684)

In line with this worldwide tendency to set up nationally integrated health information systems, health decision-makers currently intend to reinforce the Belgian health information system with a more coherent vision on longitudinal data in order to improve the monitoring of the health of its population.

A project serving this purpose is eHealth, representing an electronic platform where all people involved in the health care system can exchange information in a secure way and controllable by various instances, keeping the privacy of the individuals involved intact. The aim of eHealth is threefold: (i) to optimize the quality and continuity of health care in Belgium and the safety of the patient; (ii) to simplify administrative formalities for all actors in the health sector; and (iii) to support the development of an evidence-based health care policy (Chambre des représentants de Belgique, 2008).

The platform might not only eliminate the burden imposed on patients and health care professionals by the enormous amount of required paperwork, and reduce health care costs. As far as research is concerned, the platform might improve public health monitoring considerably by facilitating scientific – and more in particular, longitudinal - research. For example, it is expected that the health care sector will increasingly be confronted with the need for chronic care. Chronic diseases require a multidisciplinary method, which asks for more communication between different actors in the health sector. This will potentially be facilitated by the eHealth platform.

A law determining the legal framework of eHealth has recently been voted[1], but the platform is not fully implemented yet. Participation to the system is optional and not compulsory for any actor. So its functioning will depend on the interest of the people involved in the health care system and the confidence users have in the platform (VBO, 2008).

To protect the privacy of the Belgian citizens, a considerable amount of safety measures will be built into the system. According to us, the most important safeguards within the eHealth framework include (i) the presence of a legal framework, (ii) adequate institutional arrangements such as control mechanisms, and (iii) technical measures such

[1] Wet houdende oprichting en organisatie van het eHealth-platform, 21 augustus 2008 (http://www.ejustice.just.fgov.be/doc/rech_n.htm)

as encryption of the data, as well as (iv) transparency and openness about the system and the actors involved, and (v) the possibility to enter or leave the system when desired. In the following, we give some concrete examples of the data protection policy of eHealth.

Firstly, important technical standards will be installed. Data will, for example, not be centralized in one databank. The eHealth platform must be seen as an inventory of references indicating - if the patient has formally agreed - where certain information can be found, without revealing information about the health status of the patient (Claes, 2008). Also, data exchanged through eHealth will always be encrypted. Furthermore, there will only be limited access to the data available through the platform. Access will be conditioned by the use of a numeric token containing 24 digits, by identification through the use of an electronic id-card and/or by electronic signature (Cols, 2009, p.79).

Secondly, in the law determining the legal framework of the project, a considerable amount of security measures are fixed. The eHealth platform is by law forbidden to violate any of the articles included in the two basic Belgian laws on the protection of its citizens' privacy, i.e. the law of 8 December 1992 concerning the protection of the privacy related to the treatment of personal data, and the law of 22 August 2002 concerning the rights of patients (Cols, 2009, p.87).

Thirdly, the eHealth platform will be supervised by various committees and experts in the health as well as the IT sector. For example, eHealth will function under the control of a surveillance committee in which representatives of the various actors in the Belgian health care sector will be seated. The members of this committee will give their advice and vote upon the policy and strategy of the eHealth platform. For example, representatives of physicians and health support services will be appointed (Cols, 2009, p.84). Also, a consultant in information security will be appointed to control the well-functioning of the technical measures taken by eHealth. Finally, the project has been approved by the Council of State and by the Belgian Data Privacy Commission. This Commission was set up by law in 1992 as an independent organism responsible for the supervision of the protection of the privacy within research using personal data (Claes, 2008). All projects in Belgium involving automatic data processing must by law be declared to the Privacy Commission. Besides, the projects must be enlisted in the public register of the Privacy Commission. This way, citizens are informed about the necessary elements concerning the protection of their rights and of the use of personal data in Belgium. For the Privacy Commission, these requirements enable them to exercise its mission of control and to treat complaints about the privacy protection within Belgian data collection projects (Cols, 2009, p.9). A committee established within the Privacy Commission will control every project using the eHealth platform to obtain data, and will determine which person obtains access to which data and under which conditions.

No one questions the importance of health monitoring and the potential usefulness of the eHealth platform to serve that purpose. Besides, different kinds of measures have been outlined in order to protect the confidentiality of data exchanged through the eHealth platform. But the initiation of the project has nevertheless evoked a wave of negative reactions and doubts on the implications of the project for the privacy of the actors

involved in the system. Concerns have been expressed about the remaining obscurities in the law concerning the project as well as the vague organisational structures being set up (Poullet, 2008, p.130). Some physicians in particular fear the violation of the confidentiality of the patient data exchanged through eHealth, while this confidentiality is the core characteristic of the relationship between physicians and their patients. Furthermore, the use of the unique national social security number as a basis for the communication between different actors in the health sector as well as the eHealth system having access to data in the Belgian National Register both have provoked anxious reactions.

2. Privacy protection within data collection projects: various approaches

Considering data collection projects such as eHealth in Belgium, questions arise as: 'Which implications do extensive data collection projects have on people's privacy?' and 'Can the consequences of such projects be properly estimated or foreseen at all?'. Here we consider five different security mechanisms in general, which can be applied to estimate or evaluate the extent of privacy protection within a data collection project.

Firstly, the discussion concerning the impact of data collection projects could be guided by means of (theoretical) philosophical or moral basic principles or on the basis of comprehensive theories on ethical issues, analyzed by scientific researchers or philosophers.

Secondly, the level of the protection of people's privacy within a project could be estimated by an evaluation of the technical barriers protecting the data available. In the literature on the 'knowledge versus privacy'-dilemma, a great amount of attention is given to technical solutions and methodologies to secure data confidentiality. In general, two basic tools exist for a responsible distribution of information, which allows satisfying data users' need for statistical information while posing little risk of disclosure of personal information: restricted data and restricted access. To restrict data means transforming data to lower disclosure risks by means of techniques such as excluding certain attributes, 'blurring' the data by grouping or adding random error, or reversibly transforming data as in encryption techniques. Restricted access to data can be imposed by administrative procedures. The conditions on access to data may depend on the type of data user. Conditions for data sharing within an organization can differ from conditions for external data users (Duncan, 2004, pp.11-12).

A third possibility to assess the impact of data collection projects is to evaluate legal safeguards, which have the tendency to lag behind changes in technology. The legal context in which data collection projects are constructed plays a crucial role in the protection of people's privacy. Ideally, legislation recognizes the need for research access and provides sanctions for improper use of data, while recognizing the impossibility of zero disclosure risks, but limiting them as much as possible. In some countries a stable, consistent national legal framework on data confidentiality is constructed. But in many countries, the legal framework consists of various state and local public health laws, often

requiring reforms since they are outdated, fragmented, inconsistent and/or incomplete. It is crucial that legislation on the protection of the confidentiality of data is well defined, and has the least possible hiatuses, theoretical problems or vagueness (Turkington, 1997, p.114; Gostin et al, 2001, p.1389).

Legislation on privacy protection of health information should cover all health care information regardless of its form (paper or electronic), location (in storage, archives or transit), or user or holder (government, provider, or private organization). Besides, effective penalties for privacy breaches should be established. And legal safeguards that protect the privacy of health care information should be based on fair information practices. Individuals should have the right to be informed about their data being used, to control the use of their data and to review and correct personal data (Gostin, 1997, p.689).

Finally, we would like to underline the possible vulnerability of legal safeguards in times when national security is in danger. Legal regulations seem to be able to protect the confidentiality of data quite well in 'normal' times or circumstances. But they are at times placed under stress, for example in times of war or crisis. Seltzer and Anderson (2001, p.498) bring up the *"possibility of legal safeguards to be set aside in times of crisis, by legislative action, by decree, or they might even be entirely ignored. In the US, both in Word War I and II, the provisions of the Census Act on the confidentiality of data were set aside by the War Powers Act adopted after the US entered war, which provided defense authorities a way of checking confidential census information on individual Japanese Americans"*.

Fourthly, the impact of data collection projects can be evaluated by means of (practical) ethical guidelines or codes about how to deal with statistics in an ethical satisfactory manner or about how to prevent a certain system of data collection to become a threat to people's privacy. These guidelines are based on moral norms and draw statistical practitioners' attention on their potential impact on the broader society, and the ethical obligations to perform their work responsibly (ASA, 1999). In general, they address different principles and specify under each principle important ethical considerations and practical directives. Examples of ethical guidelines at the international level are the 'Declaration on Professional Ethics' by ISI (1985), the 'Ethical guidelines for Statistical Practice' by ASA (1989), the 'Fundamental Principles of Official Statistics' by the UN (1994), and the 'European Statistics Code of Practice' by Eurostat (2005), at the national level, the 'Code de Déontologie Statistique' in France by l'Association des Administrateurs de l'INSEE and les Associations des Statisticiens Economistes Anciens Elèves de l'ENSAE (1986), and the 'Code of Conduct' in the United Kingdom by the Royal Statistical Society (1993). It is of great importance that all persons treating confidential data are educated on these guidelines and are aware of the ethical assumptions and directives in the field of social research and statistical work.

Finally, organizational and operational measures play an important role in the protection of people's privacy. Examples of operational measures are: the presence of a staff with a high level of professionalism in statistical work, the distribution of information to the

public about the purpose of the project and limitations of the information provided, and revealing the methodologies used. An example of an organizational safeguard is the independence of the institutes in charge of the production of statistics, thus securing the impartiality and high quality of the statistics. Furthermore, it allows statistical organizations to resist political pressure, even in times of crisis.

3. Framework on short- and long-term implications of an extensive health monitoring

In what follows, we present a model which can serve as a framework for the evaluation of data collecting projects, complementary to the approaches discussed above. We consider some ethical implications of an extensive data collection with the purpose of monitoring health. We claim that the dimension of time is crucial in evaluating such projects: we call attention to implications both in the short and in the long run. In the short run, we consider the dilemma between an extensive data collection, indispensable for an accurate health policy, and a satisfactory protection of personal data. But long-term risks are often overlooked. We question the impact of a detailed health monitoring on future health policy regulations on the one hand and on the democratic use of data on the other hand.

	Health monitoring	
Short term dilemma	1. Health policy	2. Privacy deficit
	↓	↓
Long term risks	3. 'Health police'?	4. Democracy deficit?

3.1. Short-term dilemma: policy versus privacy

We start by looking at a short term dilemma which health monitoring projects are faced with (blocks 1 and 2 in model). On the one hand, collecting data on the population health status is crucial in addressing the increasingly complex health issues modern societies are confronted with. The accumulation of personal data within an increasingly sophisticated and automated public health information infrastructure have clearly lead to a better monitoring of the population and consequently to more accurate policies and significant health benefits and a higher quality of life. In demographic terms, a continued progression of the average life expectancy and, most importantly, of life expectancy in good health is a fact in Western societies. Although, the new technological opportunities and increasingly detailed monitoring of people's lives and health behavior contain elements that could threaten - or be perceived as threatening to - private life in society.

Citizens in modern societies might start to worry about being controlled or, 'big brother'-wise, being 'watched'. Let us take the example of the PGP-10 (Personal Genome Project), which shows a cohort of entrepreneurs and scientists who put their medical records, traits and genetic codes on the Internet. The leader of the project, George Church, hopes to get 100,000 individuals willing to have their DNA publicly searched. The aim of the project is to create a huge public database to speed up research on the causes and cures for genetic maladies (International Herald Tribune, 2008). There is no doubt that participation in this research project will serve the common good. But, one might wonder, 'What about participants' privacy?' and, since DNA is a family matter, 'What about the privacy of their family and the idea that scientists involved in the project are actually sharing information on their sister, mother, children and, even unborn, grandchildren?

When we consider it on the short term, information privacy is generally defined as the right to be left alone and to be free from surveillance and intrusion. It involves the right to control information about oneself. "*Information privacy*", according to Duncan, Jabine and de Wolf (1993, p.22), "*encompasses an individual's freedom from excessive intrusion in the quest for information and an individual's ability to choose the extent and circumstances under which his or her beliefs, behaviors, opinions, and attitudes will be shared with or withheld from others*".

The literature offers several justifications for protecting individuals' privacy right. One justification resides in the principle of respect for autonomy. Gostin (1997, p.686) explains this as *"to respect the privacy of others is to respect their autonomous wishes not to be observed or have information about themselves released*". Though, privacy rights are not only about respect for individuals' autonomy. A (perceived) violation of privacy rights might also have various practical consequences. Patients might for example develop a privacy-protective behavior towards their doctors to safeguard what they consider to be potentially harmful health information. They might 'doctor-hop' to avoid having all of their health information entrusted to one provider, withhold information, lie, or even avoid care completely. The consequences of such behavior are significant, not only for the patients and their doctors, but also for researchers and policymakers not disposing any longer of accurate health information (Goldman, 1998, pp.48-51). Moreover, as far as research is concerned, the fact that people would feel as if their privacy rights are threatened might have consequences on their research participation. For example, if potential survey participants observe instances of disclosure or perceive that confidentiality is becoming less secure, it may become more difficult for data producing agencies to obtain their cooperation (Mackie & Bradburn, 2000, p.11). Finally, economic, social or psychological damage could result from unwanted disclosure of personal information in the health domain. This way, privacy also has an instrumental value, as it permits physicians and patients to communicate more effectively, and prevents economic harm, as well as personal embarrassment and social discrimination to happen (Gostin, 1997, p.686).

The fact that privacy is a fundamental right of every citizen seems to be a widespread and universally accepted idea. But the way in which that right can and must be protected in

practice, and how far societies must go to implement this right, seems to be everything but a defined case. Too much emphasis on people's privacy rights might cause less valuable information to be available for research and other public health activities. But a more intrusive data collection might make people feel uncomfortable, lose their trust in the statistical as well as the medical system, and make patients unwilling to accurately and honestly share personal information. That is why, given the potential benefits derived from new research possibilities, it is now crucial to find the right balance between the public's right to privacy and society's need to know. To find this balance entails one of the greatest ethical dilemmas in health research nowadays. The health sector specifically is faced with this dilemma, since research in this sector concerns data reflecting some of the most personal and sensitive aspects of individuals' lives.

3.2. Long-term risks of an extensive health monitoring

The debate concerning confidentiality of data has thus far been considered from a limited point of view, by mainly focusing on the right of individuals to be protected from anyone else being informed about their personal life. However, we believe that long term implications of the way in which statistics, the collection of personal data and administrative systems are organized, should receive more attention (block 4). Collected data could in the long run for example be used for any kind of repression or entail a democratic deficit in societies, which is often not taken into consideration. Given the growing ability of our information systems to capture reality into data though, it is of the utmost importance that these data are protected properly and cannot in the long run be used for political or union repression nor for any kind of discrimination or stigmatization based on race, religion, language or culture, most likely to occur in times of political crisis. In the context of health research, sensitive personal data such as someone's sexual orientation, ethnicity or cultural background might be particularly relevant, but it is exactly this kind of personal information that is often used to base discrimination arguments on.

In what follows, we first reflect on this long-term 'privacy' risk, focusing especially on the potential violation of human rights and the possible drawback of the democracy model. After that, we focus on the long term 'policy' impact of detailed monitoring and knowledge of population health status.

3.2.1. An increased risk of a democratic deficit?

The aim of statistics is to capture in a structured way the obtained knowledge on the world surrounding us. Population data systems and statistics uncover social conditions and present them as statistical descriptions: proportion of the population below the poverty line, incidence of child abuse, extent of structural unemployment, the gender gap in similar occupations etc. Consequently, politically unnoticed social conditions might be transformed into visible statistics, which might put issues on the political agenda that would otherwise be ignored or overlooked (Prewitt, 1985, p.122).

In modern societies, statistics have become an essential element of public action. Without trustworthy statistics, all kinds of assertions (even digitized ones) might be unverifiable propagated, policies could not be evaluated, and a society would not be capable to orient itself (Deboosere and Masuy-Stroobant, 2006, p.3). That is why reliable statistics are indispensable to the democratic functioning and the cohesion of society. Objective and impartial information does only allow public authorities, policymakers and economic actors to take informed decisions (Commission of the European Communities, 2005). Statistics also facilitate an open debate about policy issues and allow citizens to evaluate the results of policies executed by public authorities. Thus, statistics are a way of sustaining and reinforcing the democratic character of public policies (Deboosere and Masuy-Stroobant, 2006, p.3).

Basically, numbers and statistics are morally and politically neutral, but what is done with them might dramatically change that assumption. Firstly, a wide range of errors associated with population-data collection might occur, potentially becoming a threat for confidentiality (Seltzer and Anderson, 2001, p.482). Secondly, possessing information about persons entails power, which can be used both positive and negative. The greater the importance of data to the securing of power, the stronger the incentives to those in power to ensure that the collected data present a favorable picture (Prewitt, 1985, p.116). Thirdly, it can be tempting to transform statistical records into administrative and surveillance records. Statistical data are used to create aggregate measures that have an impact on individuals only through significant group membership. Administrative data are collected and used to have a direct impact on individuals (Duncan, 2004, p.11). It is highly important to make a clear distinction between administrative and statistical data. Administrative use of statistical information is disastrous for the public's perception of the impartial collection of data and production of statistics. Moreover, blurring the borders creates a concentration of power and potential abuse.

Seltzer and Anderson (2001) give examples of statistical data and data systems having been used in the past to assist in planning and carrying out human rights' abuses. One of the examples given by the authors is Rwanda, where the extensive population registration system has been a tool of colonial administration during the twentieth century. Statistical reports providing the population size and its basic demographic characteristics, classified by ethnicity living in each local administrative area, together with lists of births, deaths, marriages, and persons entering and leaving the area, were used to plan and assist in the implementation of the killing operations during the genocide in 1994 (Seltzer and Anderson, 2001, p.493).

Prewitts (1985, p.126) draws attention to the potential threat to democracy, when data and statistics are used in an improper way. He mentions that: "*statistical description can bring social conditions to public attention, mobilize disadvantaged groups, and broaden the political agenda in ways that lessen the bias inherent in an electoral representation system based largely on the resources of wealth and political organization. But statistics might as well introduce practices and policies inconsistent with our traditional understanding of democracy: the objectification of politics, the assumption that what is*

not counted is not there, the temptation to substitute group membership for individual merit or need as the basis for public policy and the allocation of legislative seats according to designed racial or ethnic criteria".

An example of an event illustrating that democracy could be threatened in the long run if the new possibilities in data collection and analysis are not properly managed, is found in the United States. In 2004, the Census Bureau recognized to have made a mistake (International Herald Tribune, 2004) by communicating statistical data to the Department of Homeland Security. The Census Bureau had provided specially tabulated population statistics on Arab-Americans, including detailed information on how many people of Arab backgrounds lived in certain ZIP codes. In principle, this assistance is legal, but civil liberties groups and Arab-Americans advocacy organizations consider the incident a dangerous breach of public trust. They compare it with the Census Bureau's compilation of similar information about Japanese-American communities when internment camps were opened during World War II, for which the Census Bureau issued a formal apology in 2000 (New York Times, 2004).

We must be aware of the reality that the democratic organization model, lying at the basis of modern societies, is not guaranteed in the long run. To protect society against a potential democratic deficit as a result of the way statistics and population data systems are managed, it is therefore extremely important that data are well protected, excluding any potential abusive, undemocratic use of information enclosed in the data collected nowadays. Starting today, we should take into account the long-term consequences of an intrusive data collection or detailed monitoring of the population, and not wait until societies are actually faced with data abuses.

Different kind of measures might all together, succeed in preventing and punishing any potential undemocratic data abuse at large scale. These measures involve legal, technological, organizational and ethical safeguards, and are referred to earlier in this text. It is clear that none of the discussed measures offer an absolute guarantee for a democratic use of data, especially in the long run. However, we do believe that all these measures together can help to suppress data abuses or violation of people's privacy and human rights, by raising the financial, personal or political costs of such misuse.

3.2.2. Towards a "health police"?

Apart from reflecting on potential breaches of collected data in the long run, we question the evolution of the purpose and content of health policy measures in modern societies, resulting out of an ever more detailed health monitoring (block 3).

In Mexico, after a considerable amount of kidnappings, wealthy and even middle-class individuals are recently willing to pay $4,000 to have transmitters implanted that can indicate their location by satellite. 'Chipping' people also started in the United States where chips were inserted in 200 Alzheimer's patients for a pilot program (International Herald Tribune, 2008). One might wonder what will be next. Chipping prisoners,

mentally ill persons or teenagers? Will health monitoring extend ever further? And to what extent will certain behaviors be prohibited or obliged by regulations, in order to prevent negative effects on population's health status or well-being?

As a result of the process of individualization, the values of respect for the individual's autonomy and freedom of choice have gradually gained importance: people want to be able to choose freely. With respect to health, the next step could be to argue that people are responsible for their own health and should be able to choose freely their own health behavior (Lindbladh et al, 1998, pp.2-3). But, on the contrary, health policies seem to have become more coercive. Since we have recently been able to monitor people's health status more easily and correctly, more accurate and specific health policy measures could be implemented, especially with regard to prevention measures. On the one hand they seem indispensable with respect to today's major health issues, but on the other hand they might be considered as an intrusion in people's private life.

The discussion on smoking illustrates quite well the difficulties in designing health policy and choosing between people's free choice and striving for the common good, and the fact that attitudes of both governments and the general public can shift considerably over time. In the latter part of the 20th century, research generated evidence that smoking was harmful to both smokers and non-smokers and that second hand smoking causes the same kind of problems as direct smoking. From that moment on, the discussion regarding the justification of smoking policies became controversial. More and more countries started to implement laws banning smoking in workplaces and/or public areas and tried to prevent non-smokers from the bad effects of passively inhaling smoke, while in the mean time limiting the 'free choice' of citizens to smoke.

Research by a Swedish government commission found that a significant part of the population thinks that those who continue to ignore the consequences of a certain pattern of unhealthy behavior should pay the bill to society for being irresponsible (Lindbladh et al, 1998, pp.2-3). Some people go even further and argue that ill persons that have cared for their health must be prioritized when it comes to health care, compared to people showing irresponsible health-related behavior. One argument why it should be a government's duty to promote positive health can be found in the Constitution of the WHO, which postulates that people have a right to positive health. It states that *"the enjoyment of the highest attainable standard of health is one of the fundamental rights of every human being without distinction of race, religion, political belief, economic or social condition"* (Calman & Downie, 2002, p.397).

Finally, we illustrate the reality of increasingly intrusive health policy measures through the example of one aspect of health monitoring in Japan.

"Under a national law that recently came into effect in Japan companies and local governments must now measure the waistlines of Japanese people between the ages of 40 and 74 as part of their annual checkups. Those exceeding government limits – 33.5 inches (0.851m) for men and 35.4 inches (0.899m) for women – and having a weight-related ailment will be given dieting guidance if after three months they do not lose

weight. To reach its goals of shrinking the overweight population by 10% over the next four years and 25% over the next seven years, the government will impose financial penalties on companies and local governments that fail to meet the specific targets" (The New York Times, 2008).

This Japanese campaign is probably one of the most ambitious campaigns ever seen to tackle obesity within a society, while Japan is not even a country known for its obese population. Does current 'health policy' risk being replaced by a 'health police' in the long run, controlling our lives in a way that will be incompatible with respect for the individual's autonomy?

4. Conclusion

The success of our societies, illustrated by the continuous growth in average life expectancy, and more importantly in healthy life expectancy, is firstly related to the organization and functioning of our societies itself. But the organization of knowledge in our societies, i.e. the way in which knowledge is spread, education is given and research is organized, has also considerably contributed to this success.

Due to technological advances in modern societies, it has and will become easier to gather health data on a population. That is why it is now more than ever important to consider the consequences of the mass collection of, what is considered very personal, health and health-related information on individuals.

Through the model presented in this paper, we have called the attention to the importance of the dimension of time in evaluating the ethical impact of data collection projects. We consider it crucial to reflect not only on short-term but also on long-term consequences of such projects.

Finding the right balance between the demand for knowledge in our societies on the one hand, and the protection of its citizens' privacy and the democratic use of this knowledge on the other hand, seems to be difficult if not impossible. There is certainly no ideal solution to overcome this dilemma.

We believe that an extensive personal data collection can be compatible in the short and in the long run with a society in which the protection of people's privacy and democratic values are perceived as fundamental. We assert that, besides technical, ethical, legal and organizational measures, a democratic control on data has the ability to protect against a potential abuse of data. After all, not information itself is dangerous for people's privacy concerns, nor for the democratic system. Moreover, a highly and accurately informed public is a fundamental condition of the democracy model. The problem arises only when a limited group of people gets access to the collected information and has the intention to use it for undemocratic purposes. If a democracy does not succeed in avoiding an uneven distribution of political power, eventually due to an uneven distribution of information, power could be accumulated and in the end become harmful to democracy itself. If, on

the contrary, the collected information is shared by many, and the analysis and use of that information is dealt with in a transparent and open way, information itself is not a problem nor is it incompatible with privacy and democracy values.

Furthermore, we think that explicitly discussing ethical issues in demographic research and stimulating the awareness concerning potential problems are very important conditions to create the most optimal situation possible of balancing conflicting interests in collecting data and protecting for a potential data abuse.

Bibliography

Alderman, H., Behrman, J.R., Kohler, H., Maluccio, J.A., & Cotts Watkins, S. (2001). Attrition in Longitudinal Household Survey Data: Some Tests for Three Developing - Country Samples. *Demographic Research, 5, 4,* pp. 79-124.

American Statistical Association (1989), *Ethical guidelines for Statistical Practice*. (http://www.tcnj.edu/~asaethic/asagui.html).

Association des Administrateurs de l'INSEE, Associations des Statisticiens Economistes Anciens Alèves de l'ENSAE (1986). *Code de Déontologie Statistique.* (http://cgtinsee.free.fr/dossiers/independance/deontologie/AIS_ASTEC-fevrier-986.pdf).

Bayer, R., & Fairchild, A. (2002). The limits of privacy: surveillance and the control of disclosure. *Health Care Anal., 10*, 19-35.

Belgisch Staatsblad (2006, March 22). Wet tot wijziging van de wet van 4 juli 1962 betreffende de openbare statistiek en van de wet van 8 augustus 1983 tot regeling van een Rijksregister van de natuurlijke personen. (http://www.ejustice.just.fgov.be/cgi/article_body.pl?numac=2006011161&caller=list&article_lang=N&row_id=1&numero=1&pub_date=2006-04-21&set3=set+character_variant+'dutch.ftl'&ddfm=03&dt=WET&language=nl&choix1=EN&choix2=EN&fromtab=+montxt&nl=n&trier=afko)

Bremner, C. (2008, September 9). French revolt over Edvige: Nicolas Sarkozy's Big Brother spy computer. *The Times online* (http://www.timesonline.co.uk/tol/news/world/europe/article4703054.ece).

Calman, K.C., & Downie, R.S. (2002). Ethical principles and ethical issues in public health. In: R. Detels, J. MwEwen, R. Beaglehole, & H. Tanaka (eds.). *Oxford Textbook of Public Health. The Scope of Public Health*. Oxford: University Press.

Chambre des représentants de Belgique (2008, June 17). *Projet de loi relatif à l'institution et à l'organisation de la plate-forme eHealth* (http://www.lachambre.be/FLWB/ pdf/ 52/1257/52K1257001.pdf).

Claes, V. (2008). L'utilisation de la plate-forme eHealth est facultative. *Le Journal du médecin, 29, 1930,* p.2.

Clemetson, L. (2004, July 2004). Homeland Security Given Data on Arab-Americans. *The New York Times*.

Cols, F. (2009). *La protection de la via privée en matière de santé publique*. ([Working paper].

Commission of the European Communities (2005, May 25). *Recommendation of the Commission on the independence, integrity and accountability of the national and Community statistical authorities* (http://www.fzs.ba/Org/EUCodeOfPractice.pdf).

Dash, E. (2005, August 8). Strong privacy laws may explain data security in Europe. *International Herald Tribune* (http://www.iht.com/articles/2005/08/07/news/data.php).

Deboosere, P. & Masuy-Stroobant, G. (2006, December 1). Réflexions sur le recours aux registres administratifs pour la production statistique et la recherche démographique. *Chaire Quetelet 2006, Louvain-la-Neuve, Belgium* (http://www.uclouvain.be/cps/ucl/doc/demo/documents/Deboosere_Stroobant.pdf).

Duncan, G.T. (2004). *Exploring the Tension Between Privacy and the Social Benefits of Governmental Databases.* Paper presented at Security, Technology, and Privacy: Shaping a 21st Century Public Information Policy, 2003, April 24-25, Georgetown University Law Center, Washington DC.

Duncan, G.T., Jabine T.B., & de Wolf V.A. (1993). *Private lives and public policies*. Washington D.C.: National Academy Press.

Duncan, G.T., & Pearson, R.W. (1991). Enhancing access to data while protecting confidentiality: prospects for the future. *Statistical Science, 6*, 219-239.

Erkul, A. (2009, April 27). Geheim agente vergeet vitale handtas. *De Morgen*.

Eurostat (2005). *European Statistics Code of Practice*. (http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47141301/VERSIONE_INGLESE_WEB.PDF)

Falk, W. (2008, September 8). Beyond the conventions. *The International Herald Tribune.*

Fellegi, I.P. (2004). Official Statistics – Pressures and Challenges. *International Statistical Review, 72, 1*, 139-155.

Goodman, E. (2008, October 25). A molecular full monty. *The International Herald Tribune.*

Goldman, J. (1998). Protecting Privacy To Improve Health Care. *Health Affairs, 17, 6,* 47-61.

Gostin, L. (1997). Health Care Information and the Protection of Personal Privacy: Ethical and Legal Considerations. *Annals of Internal Medicine, 127, 8,* 683-690.

Gostin, L., Hodge, J.G., & Valdiserri, R.O. (2001). Informational Privacy and the Public's Health. *American Journal of Public Health*, *91, 9*, 1388-1392.

Goodman, E. (2008, October 25). A molecular full monty. *International Herald Tribune.*

Hodge, J.G. (2003). Health Information Privacy and Public Health. *Journal of Law, Medicine & Ethics*, *31, 4*, 663-671.

International Herald Tribune (2008, September 10). *Security plan angers many in France*.

International Statistical Institute (1985). *Declaration on Professional Ethics*. (http://isi.cbs.nl/ethics.htm).

Krishna, R., Kelleher, K. & Stahlberg, E. (2007). Patient Confidentiality in the Research Use of Clinical Medical Databases. *American Journal of Public Health, 97, 4*, 654-658.

Lindbladh, E., Lytkkens, C.H., Hanson, B.S., & Ostergren, P.O. (1998). Equity is out of fashion? An essay on autonomy and health policy in the individualized society. *Social science and medicine, 46, 8*, 1017-1025.

Mackie, C, & Bradburn, N. (2000). *Improving Access to and Confidentiality of Research Data. Report of a Workshop.* Washington, D.C.: National Academy Press.

McGraw, D, Dempsey, J.X., Harris, L. & Goldman, J. (2009). Privacy As An Enabler, Not An Impediment: Building Trust Into Health Information Exchange. *Health Affairs, 28, 2,* 416-427.

Meeus, R. (2008, December 15). Zelfs Digipass niet meer veilig voor hackers. *De Morgen*.

Myers, J, Frieden, T.R., Bherwani, K.M., & Henning, K.J. (2008). Privacy and Public Health at Risk: Public Health Confidentiality in the Digital Age. *American Journal of Public Health, 98, 5*, 793-801.

OECD (2009). Society at glance 2009 – Social indicators. (www.oecd.org/els/social /indicators/SAG).

Onishi, N. (2008, June 13). Japan, Seeking Trim Waists, Measures Millions. *The New York Times*.

Poullet, Y. (2008). Construire un cadre juridique pour l'e-Health. In: J. Herveg (ed.) *La protection des données médicales. Les défus du XXIe siècle*, 89-130. Louvain-la-Neuve: Anthemis s.a.

Prewitt, K. (1985). Public statistics and democratic politics. In J.J. Smelser and D.R. Gerstein (eds.). *Behavioral and Social Science: Fifty years of Discovery*. Washington, D.C.: National Academic Press.

Rossenbach, I. (2008, July 15). Un passeport biométrique que ne passe pas. *Rue 89*. (http://www.rue89.com/2008/07/15/un-passeport-biometrique-qui-ne-passe-pas).

Royal Statistical Society (1993). *Code of Conduct.* (http://www.rss.org.uk/pdf/Prof%20memb%20-%20code%20of%20conduct%20new%20charter.pdf).

Seltzer, W. (2001). *U.S. Federal Statistics and Statistical Ethics: The Role of the American Statistical Association's Ethical Guidelines for Statistical Practice* (http://www.uwm.edu/~margo/govstat/wss.pdf).

Seltzer, W. & Anderson, M. (2001). The Dark Side of Numbers: The Role of Population Data Systems in Human Rights Abuses. *Social research, 68, 2*, 481-513.

Seltzer, W. & Anderson, M. (2002). *NCES and the Patriot Act: An early appraisal of facts and issues.* Prepared for presentation at the Joint Statistical Meetings, New York City, August 12, 2002.

Timmerman, G. (2008, October 23). Megadatabank voor politie in de maak. *De Morgen.*

Turkington, R.C. (1997). Medical Record Confidentiality Law, Scientific Research, and Data Collection in the Information Age. *Journal of Law, Medicine & Ethics, 25*, 113-129.

United Nations Statistics Division (1994). *Fundamental Principles of Official Statistics*. (http://unstats.un.org/unsd/methods/statorg/FP-English.htm

Verbond van Belgische Ondernemingen (2008, December). Overheidsmanager Frank Robben lanceert het eHealth-platform. *Forward* (http://www.law.kuleuven.be/icri/frobben/documents/Forward%20-%20december %202008%20nl.pdf).

Yee, J.L., & Niemeier, D. (1996). Advantages and disadvantages: longitudinal vs. repeated cross-section surveys. [Discussion paper]. (http://ntl.bts.gov/lib/6000/6900/6910/bat.pdf).