

# **Does research threatens privacy or does privacy threatens research?**



*Anonymisation of personal information*

**Master thesis**

**Evelien Y. Van Roijen**

**Supervisor: Dr. Marcel Verweij**

**Applied Ethics**

**University of Utrecht**

**August 2011**

|  |    |
|--|----|
| Abstract.....  | 3  |
| Introduction .....   | 4  |
| 1. The practice: data gathering by a GP used by third parties .....                    | 6  |
| 1.1 Applicable laws and regulations.....   | 7  |
| 1.2 Analysing legislation .....  | 8  |
| 1.3 Dutch future steps for accessibility to medical records .....                      | 9  |
| 2. Research on genetic information .....   | 10 |
| 2.1 Examining current regulations related to health databases.....                     | 11 |
| 2.2 Genetic exceptionalism.....  | 12 |
| 3. Data mining and the risks for infringements on privacy .....                        | 14 |
| 4. Protection of personal data .....   | 15 |
| 4.1 Confidentiality .....  | 15 |
| 4.2 Having information differs from having knowledge .....                             | 18 |
| 4.3 Privacy interests.....   | 19 |
| 4.4 Personal information.....  | 20 |
| 4.5 Group interests.....   | 20 |
| 5. Informed consent and its inadequacy for protecting privacy in health databases..... | 24 |
| 5.1 Morally justified research without informed consent .....                          | 26 |
| 5.2 Additional moral safeguards.....   | 28 |
| 5.2.1 Trusting the research community .....  | 28 |
| 5.2.2. Rules for the use of information by researchers .....                           | 29 |
| 5.2.3. Warrant assent .....  | 30 |
| 5.2.4 Conditions for warrant assent.....   | 31 |
| 6. Conclusion.....   | 33 |
| Literature .....   | 36 |

## Abstract

From 1970 onwards the General Practitioner (GP) replaced the handwritten medical cards by Electronic Health Records (EHRs). While at first the information was stored on the GP's local server, in the 90-ies computers were linked and gradually more persons had access to EHRs. While currently the EHRs are on Local Health Servers in the region, advocates in health care strive to have them nation-wide accessible to them. The EHRs, however, contain sensitive information and in the Netherlands the Law on protection of personal information (WBP) regulates data protection, including sensitive information. Sensitive information however is hard to define. So part of the WBP is based on the control of patients over the content of EHRs and the obligations researchers have to use anonymised records or ask data subjects for an informed consent. Traditionally, information on EHRs relates to the past, but nowadays we have the tools to make predictions about risks for future diseases by looking at a persons' genetic make-up. A persons' genetic information is unique, so the source of biospecimens can theoretically always be traced. This challenges the privacy protection of data subjects providing researchers access to their HER, even anonymised. Therefore I scrutinise in this paper which moral safeguards are to be taken into account to protect research participants privacy when they provide their biospecimens to a national health database for general research. Adding to our moral concerns is the technique of data mining. Data mining is a technological process searching for correlations or patterns in persons' data. The outcome is a group profile, e.g. persons driving red cars are likely to contract cancer. Life insurance companies may adjust their premiums based on this group profile. The public may then rightfully make a moral appeal about protection of privacy and confidentiality. However, as argued in this paper the traditional concept of direct bilateral confidentiality in a physician-patient relation has yield with the linkage of computers. We therefore need to adjust our concept of confidentiality. Although more people have access to EHRs, so more confidential information is available to more persons, this does not mean that these persons also have knowledge. The significance knowledge has for a person depends largely on the data holder's cognitive and practical commitments. Information acquires normative significance only because it can be used in certain actions and therefore possession of information is significant, only against some background, while we tend to believe that information itself has intrinsic value. So, instead of the focus of the WBP on right to informational privacy over content, it might be more enlightening to change our focus to the use of personal information. The justification for informed consent has primarily been data subject's autonomous choice, manifested in balancing risks and benefits. However, although informed consent is important, the foundation can not be found in data subject's autonomy, since there is no specific proposal to consent to by providing your biospecimens to national health databases. I therefore propose therefore to use the notion of "warrant assent" as a different way of thinking how individuals can confirm their willingness to have their biospecimens used for general research questions. Assenting implies a supporting attitude of the person to a more general practice or institution. However, data subjects need some warrants besides trusting the research community and I outline eight suggestions, which might be helpful for data subjects to feel confident enough to provide our biospecimens to national health databases and researchers aware that they need to act respectful and responsible.

## Introduction

From the 1970s onwards the handwritten “medical green cards” in the Netherlands were gradually replaced by computer files. While at the beginning these files were only available on the local computer of the General Practitioner (GP), pretty soon these data were shared on Local Health Servers with multiple persons having access with the aim to provide adequate health care in cases of emergency or if another physician replaces the regular GP. It took only two decades before one out of every twenty-five citizen in the Netherlands had authorised access to medical files of individuals. These persons are all in one way or the other related to the health care system. The traditional relationship between practitioner and patient has therefore stretched to increasingly more persons having access to someone’s sensitive and confidential information about health, socio-economic status, family histories of diseases, risk markers for e.g. obesity, heart stroke or alcohol addictions. This information is obviously interesting for more parties than just health care providers. Life insurance companies, employers, lenders and possibly family members and partners might also like to have full access to this information, although for quite different reasons. Due to this increasing number of persons having access to medical records, protecting medical health privacy became a hot debated topic in the last five decades.

Besides the extension of access to health information of individuals, other kind of health information about individuals came available which some person might consider to be even more problematic if known to third parties. In the early fifties of the last century two researchers in the UK, Rosalind Franklin and Maurice Wilkins, discovered structures of DNA, which information were used by James Watson and Francis Crick to build a model to show how our double helix looks like. On 28 February 1953 Watson and Crick felt they had solved the problem enough for Crick to proclaim in the local pub the Eagle in Cambridge, that they had “found the secret of life”<sup>1</sup>. It took researchers years to improve the model and interpret the data, but eventually in 1990 a project<sup>2</sup> started in the United States to identify all of the estimated 20.000-25.000 genes in human DNA. This project intended to be finished in 2005, but came to a conclusion already in 2003. Ever since the project started and information became available, genetic information became increasingly an integral part of the corpus of health data. Using many different biospecimens contributed greatly to the advance of the research on genetics and international collaboration is therefore strongly promoted<sup>3</sup>. Researchers started to look for the genetic underpinnings of a disease, even if little is known about the underlying cause. To be able to do this research requires large numbers of tissue samples either from families in whom heritable diseases are prevalent or from diseased tissue samples and normal tissue samples to distinguish variations (the so-called Genome Wide Association Studies – GWAS). For that reason general national health databases has been set-up in different countries<sup>4</sup>. These large collection of identifiable samples have been deemed indispensable not only for (epidemiological) biomedical research, but also for forensic needs, autopsy requests, detection and prosecution of criminals and in cases of parental and child disputes. Although bio-medical research is of tremendous value for

---

<sup>1</sup> <http://www.youtube.com/watch?v=sf0YXnAFBs8> (last visited 04-08-11)

<sup>2</sup> [http://www.ornl.gov/sci/techresources/Human\\_Genome/home.shtml](http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml) (last visited 04-08-11)

<sup>3</sup> <http://www.p3observatory.org/> (last visited 04-08-11)

<sup>4</sup> E.g. Austria, Estonia, Iceland, Norway, Sweden and U.K.

large parts of the populations, it needs to be balanced against privacy concerns of individuals, if identifiable genetic information can become available to e.g. insurance companies, employers or financial institutions. Because of this threat of privacy invasion and the possibility of discrimination and stigmatisation, an appeal has been made by several authors to treat genetic information differently than regular medical records and provide genetic information an exceptional status by granting this information a higher level of protection than regular medical records.

When people can be harmed through infringements of their privacy by giving their biospecimens to health databases or tell the GP in confidence their secrets, there seem to be one simple solution: “Do not do that!” But this can create practical and severe problems, which could easily backfire to the individual. Firstly, not telling the GP the whole story or even avoiding her might seriously jeopardise a persons’ health. Most likely they will not receive the treatment they need, which eventually even could cause early death. And secondly, if no-one is willing to provide biospecimens to health databases for research, there will be no research. We highly value health and improvements of health, prevention of diseases, better diagnostics and treatments of diseases are on top of our wish list. If we can prevent others from being harmed to provide our biospecimens for research and this is with only minor efforts and without much risk, it would be morally wrong if we restrain from this. Although risk may be sometimes absent or minor, there is also a possibility of high risks of infringements of privacy when we are direct or indirect identifiable to others in publications or if we find out information about ourselves which we did not want to know. How can we avoid potential privacy risks factors? Although technical precautions in databases should be well in place, reflected on and evaluated time and again, especially with data-mining and multi-institutional linked computers, I am interested in the moral dimension of the privacy risk factors. Therefore this study aims answering the question: what should morally be taken into account to protect privacy interests of individuals when they provide human tissue to health databases for bio-medical research? Therefore, in the first chapter I will outline non-genetic medical information practices and regulation and provide an analysis of possible issues of concerns, which research participants might have if they provide their biospecimens to a health database. Since the biospecimens contain genetic information, which is gathered as being more sensitive than medical records, I will scrutinise in the next chapter if this is the case and if there are moral grounds to justify exceptions related to genetic information. Since technologic developments are up to full speed, new privacy risks are added with the emerging linking of computers and databases, so I will therefore briefly explain in chapter three the risks involved for privacy infringements in a rather new computer technique called datamining. Data mining is the application of specific algorithms in databases to ‘discover’ and extract patterns or correlations in data in stead of looking at causal relations. In chapter four I will establish if our assumption that we are safeguarded by the moral concept of confidentiality in a direct physician-patient relation is correct. In this chapter I will narrow privacy protection down to informational privacy to scrutinise the privacy risks people have on an individual bases, but also as identified member of a group constructed by datamining. If mandatory informed consent requirements are sufficient to safeguard privacy interests of individuals and groups, will be the topic of

chapter 5. I will finalise this examination with some suggestions for moral safeguards to trust health databases enough to provide our biospecimens for medical research for the public good.

### **1. The practice: data gathering by a GP used by third parties**

In the Dutch health care system general practitioners (GP) keep health records from birth till death of almost every citizen in the Netherlands, as a cradle to grave card, which is more or less the same in the rest of the Western world. The information gathered is not only medical information, but also genealogical, social and in some sense financial information as the socio-economic circumstances. The GP starts the file when a baby is born, so most of the time it is this family doctor who gathers the first information and keep the record up to date, if no other health problems appear. Sometimes the GP is part of a larger GP practice and if he or she is not on duty, colleagues have access to the file in case of emergencies. The aim to keep a well-documented digital medical record by the GP is to provide the individual patient the best possible care not only by herself, but also by other health care providers, if necessary. So, if a patient has to go to a hospital for treatment, specialists, nurses and other care providers have access to most and sometimes all of the information in the medical record of the individual. The advantages of sharing medical information amongst health care providers are numerous for the patients: the patient does not have to tell the whole story again, most likely to forget something essential like an allergy to medication; results from previous tests are available in the record and can either be checked with new tests or used to start treatment; current medication is mentioned in the file, so a specialist can notice if there will be a conflict between old and new medication. So, due to electronic developments and the implementation of the Local Health Server at the end of the last century more sensitive health information became available with more people having access to this information. In 1997 already 120 municipalities had multidisciplinary partnerships sharing medical records of over 4 million persons. 700 general practitioners, 400 pharmacists, 40 health care centres were sharing more than 10 million recipes every year<sup>5</sup>.

Besides the health care providers also health insurance companies and outsourced administration offices have access to parts of the medical information of individuals, since they handle the requests for reimbursements of costs for medication and treatments. They also receive the diagnose code once a patient leaves the hospital, so they are aware of upcoming costs and persons who might become a financial risk. Health insurance companies are most of the time for profit companies and in an ideal world they would like to have as much profit as possible every year. To attain their profits, they try to avoid as many persons as possible bearing a financial risk. However, they are not free to reject persons access to a health insurance, because due to the Dutch health care system, insurance companies are obliged to accept everyone, although they are allowed to offer them only a basic insurance and charge higher premiums for persons at risk.

And besides insurance companies, researchers as third parties have access to individual medical records as well. Before the eighties of the last century, it was general practice that researchers had access to full medical

---

<sup>5</sup> Spaans JAJ, Hoonakker DGPH, Leer van der OFC (red), 1997, p. 83

records without any consenting from patients allowing them to do so. Most of the time patients were not aware of this practice and since it was common use in the research practices, researchers gave hardly any attention to the ethical implications of their actions. But, also research techniques evolved and more and better research is possible and done to improve the health of the public, find better treatments and drugs and better diagnostic methods. So, researchers are now even more interested than e.g. a couple of decades ago in medical records to do their research, whether it is for a research on a heritable disease or to compare their findings with the normal standard, persons not having these diseases.

To sum up, because of the quick evolvments of the digital highway in the last five decades, more and more people have now access to more and more sensitive information in medical records of individuals. This has implications for the privacy of these individuals. The international community and the Dutch government responded to these developments and brought new legislation into force protecting people against invasion on their private lives. I will now first examine if and if so, how current legislation covers the protection of sensitive information of individuals in medical records.

### 1.1 Applicable laws and regulations<sup>6</sup>

Wanting to use the upcoming digital highway to collect and store all personal data of their citizens under a personal number, the Dutch government decided to have a new census in 1971. The public objected, because it was unclear to them why and by whom the information was going to be used and for what reasons. This moment became the starting point for a debate about privacy rules and regulations. The research community stepped in into the debate, addressing that when privacy regulations became too strict, scientific research would be difficult to do, if not impossible. The result of the debate running up to the final legislation of the WPR (Wet Persoon Registratie – Law Registration of Persons), which came into force on July 1<sup>st</sup>, 1989, was that this law was much more about self-regulation of the different professions than normative rules about what is forbidden and permitted.

At the same time the international community reacted on the evolution of technology and the increase of international cooperation and trade and the Council of Europe came up with a Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data in January 1981, which was based on the OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data from 1980. The Netherlands ratified the Treaty of Strasbourg in 1990 and although it is not strictly speaking a law, it is a strong political guideline for the Dutch government, being a party, to shape Dutch regulations along these lines. The Treaty however, left a lot open for interpretation and the European Commission noticing that lack of harmonisation in privacy regulation among its member states became an obstacle for the internal market and came with a general data protection directive in 1995 (95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data) and a more specific directive in 2002 (2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic

---

<sup>6</sup> This chapter is based on chapter 3 of the dissertation of M.C. Ploem, 2004

communications sector). Since the Dutch government needed to implement the EC Directive, they updated the WPR and a new legislation came into force in September 2001, the WBP (Wet Bescherming Persoonsgegevens - Law protecting personal information).

In the Netherlands the KNAW (Koninklijke Nederlandse Academie van Wetenschappen – Royal Academy of Science) and the NWO (Nederlandse Organisatie voor Wetenschappelijk Onderzoek – Dutch Organisation of Scientific Research) are members of the European Science Foundation (ESF). On November 12<sup>th</sup>, 1980 the Assembly of the ESF accepted a declaration, notifying the tension between privacy interests of individuals and the research interests in personal data. Before this date the research community was never held accountable and could work freely with all information available to them without patients knowing or consenting to this. But the Foundation also noticed that technological developments and increased cooperation could raise conflicts between their need for the gathering, storage and use of data and personal privacy concerns. The codes dated from 1980 had already been renewed in 1985. The main reason was that epidemiological and longitudinal research ran into obstacles doing their research with identifiable information if they committed to this code. The research community agreed that personal information should be anonymised as soon as possible and if possible, but longitudinal and epidemiological research by definition need access to identifiable information.

## 1.2 Analysing legislation

Looking at the current regulations Ploem noticed a shift with the previous regulations. In the new legislation the interests of research can outweigh privacy interest of individuals. The interests of the research community are facilitated by a smoother regime for using personal information for research and statistics. Personal information is only very generally defined, mainly because the WBP is intended to be a general, sector and technology independent legislation<sup>7</sup>. The underlying assumption is that use of personal information for epidemiological research is less intrusive than using it for administrative reasons. In case of the latter the information is used to make decisions on an individual basis, which could be harmful. Researchers are by no mean interested in the individual, they are interested in the aggregated data of a population. The facilitating norms are divided into firstly, a smoothen up regulations regarding the collection of information, i.e. availability and accessibility of the information and secondly smoothen up the further storage and use of the collected information.

Regarding the collection of data, researchers need to have a specific aim to collect, and cannot collect more information than necessary for their research question. However, they are allowed to expand their search for information to entire EHRs, when their research question is of general interests. Improving public health is the most common term to collect as much information as possible. Researchers have no duty to report their research anymore if already collected data is anonymised. Informing the subject that his or her data is used in new research has also deemed not to be necessary if this implies a disproportionate effort for researchers to trace the data subject. And although privacy has never been seen as an absolute value, it is now clear that with the coding techniques where information becomes anonymous the researcher has less obstacles to continue her research and aggregate and store data, while it also can be kept longer and used for further research questions. So, according to

---

<sup>7</sup> First Phase evaluation WBP (Eerste fase evaluatie Wet bescherming persoonsgegevens), p.166



the WBP, identifiable sensitive personal information data is available for secondary use, although this secondary research must have a relation to the previous research question. However, a general collection, storage and use of biospecimens are not allowed according to the WBP. With the current WBP it would be impermissible to start a national health database in the Netherlands. Researchers can get around by setting up a specific database<sup>8</sup>, related to one disease, e.g. cancer or diabetes or a specific group, e.g. family members with a hereditary disease or twin cohorts, anonymised the data and if their research question is of general interests and deemed necessary it is permissible to do so.

### 1.3 Dutch future steps for accessibility to medical records

A next step aimed by Dutch politicians is to allow regional networks, the Local Health Servers, to expand into a national infrastructure where all authorised caregivers have access to the medical records of every individual. Authorised care givers are registered persons according to the Law on the Professions in Individual Healthcare Systems (Wet BIG - Wet op de Beroepen in de Individuele Gezondheidszorg - 1993). If all caregivers, registered according to Wet BIG, half a million people will eventually have access to all medical records, which is 1 in every 25 citizen and it is likely that someone in your social network is one of them. However, not all caregivers will get the same kind of authorisation to access the medical records, this will depend on the authorisation given to them, which depends on their status as health care provider. A physician has access to more information than a nurse for instance. Both the EC Directive and the Dutch WBP gave room to self regulation of the professions within the framework set out. The core concept of the law, avoiding breaches of privacy, are considered to be resolved or at least contravened by the use of unidentifiable datasets according to the science community. Since all Dutch medical records are potentially available to researchers, breaches of privacy will remain possible. This goes further than just medical information, since patients tend to give their physician in confidence more information, which might include relational information, family history, sexual behaviour and sometimes financial information. General practitioners write down extra information in the records, which might contain value-laden remarks. The fact that confidential information given in a physician-patient relation can be seen and used by many others might mean that we need to rethink the concept of confidentiality in medical settings. I will get back on this later on.

Nothing is said in the Dutch law on Medical Treatment Agreement (Wet op geneeskundige behandelingsovereenkomst – WGBO, which came into force in 1995) about the access of medical researchers to medical records, mainly because this law is directed to health caregivers. We might assume that the researchers are provided access only after approval by a medical ethical review board to do research on certain patients. There is however, no one single login code for researchers, to the Local Health Server and this will also not be the case in the national database (EPD – Elektronisch Patiënten dossier - Electronic Health Record -EHR), since researchers are investigating totally different kind of research questions and their need for information is completely depending on that specific question. For that reason researchers cannot be treated as a category, while

---

<sup>8</sup> The Dutch Pearl Necklace Initiative (het Parelsnoer initiatief) is an example of biobanks related to eight specific diseases

nurses, dentists and physicians can. In practice it is very likely that the treating physician of the patient in a hospital, who most of the time is also the supervisor of the researcher provide their own login code to the researcher for additional information about the patient, in which case all information in the file is visible to the researcher. So, if and when this national infrastructure will be put in place, privacy concern of individuals will only increase, because the researchers have access to all persons in the Netherlands instead of only to those who fall under the treatment of their supervisor, which is the physician-investigator. But still, without the national infrastructure in place, because the Dutch parliament rejected the proposal, individuals already concern about privacy infringements on a regional level.

## **2. Research on genetic information**

Since the accomplishment of the Human Genome Project (HGP) in 2003 a large part of biomedical research shifted its focus from phenotype to genotype. From the approximately 20.000 – 25.000 genes more than 6000 genes are identified as single-genes or Mendelian disorders and are inheritable in different kind of patterns. Some grave examples of single-gene disorders are cystic fibrosis, sickle cell anaemia and Huntington's disease. Most of the diseases known so far are far more complex with multiple genes playing a role one way or the other. To be able to examine the structure and functions of genes large databases are necessary to compare genes. Since the function and structure of genes also differs slightly between different populations, it is helpful for researchers to have a nation-wide population based bank storing all kinds of human biospecimens for examination and comparison. Society has an interest in better medication with fewer side effects, quicker and better diagnosis and therefore accepts and approves biomedical research on human biospecimens within certain limits. One of the limits is related to the fact that genes are uniquely identifiable including identical twins (although in that case based on the so-called SNPs), who have the exact same set of genes. So, while medical records stripped of all personal identifiers might provide some certainty to avoid breaches of privacy, genetic information from human biospecimens stripped of all identifiers can easily be matched with the "owner". While sequencing of genetic information in general is in contrast with medical information neutral, the analyses and interpretation of this information can have major impact on individual's life. If genetic information mentions that the ATM, p53, CHEK2, PTEN, or CDH1 genes are present, it does not mean anything unless it is analysed and interpreted as being abnormal, enlarging the opportunity to contract cancer. Therefore genetic information can be seen as sensitive information, because it can be analysed further and we have interests protecting it. How does research deal with our genetic make-up in practice?

Most research institutes have their own storage of human biospecimens already for decades in place. Some are small of size, having one refrigerator containing blood samples, others may be large in scale, containing thousand of blood samples and tissues in all kinds of refrigerators or cryogenic storage, which means the biospecimens are kept in liquid nitrogen with a temperature of  $-196^{\circ}\text{C}$ .

Currently, different terms for health databases are used interchangeable. I found the words bio bank, tissue bank, gene bank, DNA bank and also human genetic research database, genomic database, tissue

establishment, DNA repository. In most cases the tissue collection and blood samples stored at one place are separated from the databases that contain information only, like medical records, medical photographs or family history. In these cases both physical and informational material is most of the time in the same building. Due to evolving technological developments it is now possible to connect computers, which are geographically distributed, and also otherwise independent, the so-called Grid computing and e-science. For researchers these are processed as one virtual health database, because they can interrogate and search as if they were one single, centrally organized entity. For instance TuBaFrost<sup>9</sup> is a central European database of information derived from frozen tumour samples stored in numerous different countries and has members and guardians all across Europe. There is not one single person or institute responsible or accountable for all data and data transfer. To address these different kind of stored, all kind of physical bio-samples and digital genetic information and the different processes, be it one computer or several linked computer plus the spreading all over the globe or just at one location, I will use the term “health database” as a working concept.

Most health data can be collected and stored in identifiable or non-identifiable form. Identifiable health information may be defined as any information, whether in oral, written, electronic, visual or other forms, that relates to an individual’s past, present and future health status and which reveals the identity of the individual whose health status is the subject of information. Conversely, health information that does not reveal (or cannot be linked with other information to reveal) an individual’s health status is often defined as non-identifiable or anonymous. However, due to continuing technological development, individuals with access to stored tissue are increasingly able to discover the identity of individual donors using genetic markers<sup>10</sup>. Also, the international dimension in research and data linking and merging can de-identify patients again. Anonymous information does not always remain without identification. Anonymous data or biospecimens does not have to remain unknown forever. It is easier to trace back the ‘owner’ of genetic information if you have a possibility to compare and match information, linking more databases will increase the risk of this to happen.

## 2.1 Examining current regulations related to health databases

In the Dutch law genetic information is not specifically addressed, nor do we find any information about this in the EC data protection directive. We address genetic information as sensitive information in which case the same legislation applies to genetic information as to medical health records. That implies that either researchers use genetic information because a donor wrote a consent form<sup>11</sup> or because an ethical review board has approved their research proposal or the genetic information is duly stripped from identifying markers. Some countries however, have applied specific legislation to genetic information, they address genetic information as personal, even

---

<sup>9</sup> [www.tubafrost.org](http://www.tubafrost.org)

<sup>10</sup> although there is no 100% guarantee, as is shown in the Tilburger murder case, where the police found a DNA match in a forensic computer, which caused them to arrest Durgale M. identified as being the murderer of Cees van der Wiel in 1999, until after eight months a witness testifies that someone else had murdered the man. The used identifier marks matched exactly, but other marks revealed the differences.

<sup>11</sup> I will not discuss any kind of tissues withdrawn during operation before the legislation came into force. This is a separate ethical problem, which will not be addressed in this paper.

unique, and even without identifiers, due to its uniqueness, and it needs special regulations. Therefore, I will first scrutinise if there is a moral justification to grant genetic information a special protection.

## 2.2 Genetic exceptionalism

For some authors genetic information is so different than medical record information that it deserves special treatment in legislation. It is claimed to be ‘special’, because it can be predictive, it is immutable, it is personally identifiable and it may have implications for others, like family members or relationships in the case of reproduction choices. The special characteristic of genes leads Annas, Glantz and Roche (1995) as part of the ethical, legal and social issues (ELSI) programme for the Human Genome Project<sup>12</sup> to categorise genetic information as unique. In their view uniqueness has at least two aspects. On the one hand genetic information is unique because it is presumed to be very precise in its ability to give information about an individual present and future health status, a characteristic that other forms of health information do not have. They state that genetic information is like a diary for potential future medical problems. Genetic data may accurately predict whether an individual will develop a specific condition like Alzheimer’s disease or contain information that may be unlocked by future scientific understanding to reveal diseases. On the other hand genetic information is unique according to Annas et al. compared to other health information in the sense that it is not information ascribable to any individual, but only to one specific individual and in some way also to family members. In their view genetic information has implication for life choices, insurance and employment and raises the spectrum of discrimination and stigmatization against individuals and groups. Genetic information also has ramifications for relatives in which case the balance can be shifted from the rights and interests of the individual to the family members and relations. Hence, these proponents of the so-called “genetic exceptionalism” see genetic material as being uniquely powerful and personal and threatening privacy interests more than medical health records and therefore require a higher level of protection. This hypothesis that genetic data are unique supports the enactment of legislation affording genetics a special status, but is genetic information really unique opposed to medical records to morally justify special protection?

Some authors (e.g. Murray, 1997; O’Neill 2002) disagree. They argue that many sources of medical information share similar characteristics and genetic data does not require a special status. Genetic information alone cannot begin to provide the level of detail of an individual’s medical and social profile as do health records generally. Non-genetic health records contain an array of personal and social information with multiple uses e.g. 1) demographic information such as sex, age, race and occupation; 2) financial information, such as income and employment status; 3) information about disabilities, special needs and other eligible criteria for government benefits and 4) medical information such as diagnosis, treatments and diseases of family history.

If we look closer to the different arguments used by the proponents of genetic exceptionalism, we might figure out whether genetic information needs to be better protected or merits different regulation than medical

---

<sup>12</sup> [http://www.ornl.gov/sci/techresources/Human\\_Genome/elsi/elsi.shtml](http://www.ornl.gov/sci/techresources/Human_Genome/elsi/elsi.shtml)

health records. <sup>13</sup>Firstly, genetic information is seen as a predictor of grave diseases. And indeed while some genetic illnesses are limited to a single gene disorder, as Tay Sachs or sickle cell anaemia, and are certain predictors, the large part of diseases, such as cancer, heart diseases, schizophrenia and depressive illness have multiple genetic components, not easily to trace and highly unlikely to predict with certainty. Besides genetic predictions, some other, non genetic indicators may predict as well with some degree of certainty whether a person will develop a certain malady. High blood pressure, elevated cholesterol, obesity, ingestion of a lot of caffeine, use of alcohol, drugs or tobacco are just a few examples.

A second suggested difference between genetic and non-genetic information is that the genetic feature of a person is fixed, unchanging and unchangeable. Mapping DNA is mapping fate. People feel stuck with their genes, while their personal health may largely be felt as under control. Individuals have the impression that illnesses can often be reversed through personal changes in environment, behaviour or clinical intervention. This view is misguided, since genetic flaws can be altered or corrected as well through clinical intervention and changes in environmental changes can compensate for genetic propensities. Genetic information has been interpreted by some authors as being more deterministic than today's research has been able to confirm, it realistically provides only a glimpse of what makes persons susceptible to disease and other conditions.

Another suggested difference is that genetic data are qualitatively different from other health information because they are inherently linked to a single person who is the subject of the information or host of the biological sample. While non-genetic, non-identifiable health records could apply to many individuals, genetic data are uniquely personal and portable. And yes, potentially any sample of tissue or DNA is traceable to a single individual, however it is not the only unique feature of the physical body. The science of biometrics focuses on the many ways individuals can be identified by their unique characteristics. Each person's fingerprint, hand or face geometry, voice spectrograms and handwriting samples are sufficiently distinctive to accurately identify individuals with the assistance of modern technology and expert methods of analyses. The human iris is arguable a better personal identifier than DNA, because each person's iris, even those of identical twins are unique.

A fourth suggested difference between genetic and other medical information is that genetic information does not simply reveal health and personal characteristic about the individual, but also about the family, their parents, siblings and children. When a woman learns that she has a high genetic propensity for breast cancer, her mother, sisters, aunts and daughters who share the same genetic code may also be at risk. But family medical histories have long been an important component of clinical practice. Mental disorders, alcoholism, heart disease and cancer as familial disorders are potential predictors of maladies in family members, so genetic information is not the only feature to focus on. Proponents of genetic exceptionalism seem to overlook that most diagnoses are made from direct observation and family history.

---

<sup>13</sup> This chapter is also based on conversations with Prof. Peter Heutink, head of department Medical Genome analyses (VUmc) and his staff

Genetic predisposition may be considered as another risk factor, just as social and environmental risks are. This conception of human aetiology makes the distinction between genetic and non-genetic information troubling. Treating genetics as distinct may discourage individuals from seeking testing and treatment and ultimately thwart future scientific progress. It also is hard to distinguish genetic from non-genetic information in an EHR. Focusing only on genetic privacy might convey the perception that the public need not to worry about the confidentiality of other kind of medical information. They would foster complacency in an area where insufficient protection may exist. Treating genetic information as exceptional might become a risk because regulation of the digital information might not end up with the same level of protection.

It is the dissemination of the information that is of ultimate concern to individuals, not just the genetic results from a test done. But, where traditional health information is increasingly exchanged in electronic form in a (inter)national health information infrastructure, the potential for multiple disclosures exists for genetic data as well. I will therefore continue using genetic information and medical health records as one health database worth protecting for privacy interests.

An added risk is the development of new technology, so called datamining, uses for its research health information as well. But instead of deduct results out of tests as traditionally is done in research, this technology is inductive. The results out of these researches can have a major impact on privacy, but remain unregulated in the WBP. I will explain this technology a bit more to include these added risks in this analyses.

### **3. Data mining and the risks for infringements on privacy**

Besides the extended accessibility of medical information a new technology is imposing an added risk to infringements of privacy. One of these developments in research settings is data mining. Data mining is the application of specific algorithms to 'discover' and extract patterns or correlations in data. There is another process called KDD (Knowledge Discovery in Databases) process, which consist of three steps. The first step is data gathering and is done before the data is searched for patterns or correlations and includes the work done after searching, namely analysis and interpretation of the data. For comfort reason I will use data mining involving the last process of KDD.

Information acquired about persons via data mining is often derived from implicit patterns in the data, to discover new facts or relationships about that person and create new groups. While observation and experimentation usually consider only causal relationships, data mining investigates a broader range of possible relations, including classifications and correlations. This technology can be used to generate group profiles on a large scale, which can be very helpful to guide policy. There is a major risk however, that since data mining is inherently inductive, many rules may be overtly generalised and appearing to be useful but are instead misleading<sup>14</sup>. Although data mining shares a great deal in common with statistics, since both strive towards discovering some structure in data, data mining also draws heavily from many other disciplines. Data mining differs from statistics

---

<sup>14</sup> Fule and Roddick, 2004, p. 159

in that it must deal with heterogeneous data fields, not just heterogeneous numbers, as is the case in statistics. Although we are used to be classified in groups and are not surprised if we find ourselves in the group profile of house owners or women, the process of data mining can discover or construct new facts about individuals, which might surprise us. An owner of a red car might be surprised to learn that s/he is assigned to a group of individuals likely to have or contract cancer, because of some arbitrary correlation. The problem for data mining researchers is that investigations using knowledge discovery tools are commonly open-ended – it is not possible to know what will be found until it is discovered. Also, many useful investigations require the use of non-anonymised data, e.g. link episodes of treatment. Vedder (2001) notes that because the kind of information contained in many group profiles being a result of a data mining process falls outside the sphere of protection granted to personal data, this is not protected by our current privacy laws. Privacy laws and data protection guidelines have been instituted to protect personal data used in earlier computerised database-retrieval techniques, where sensitive data is 1) explicit in a database, 2) confidential in nature and 3) exchanged between or across databases. In data mining processes personal and general information is used and these normative protections however do not apply to personal data manipulated in data mining processes, where personal information is often implicit in the data, non-confidential in nature and not exchanged between databases. We may conclude that individuals rights to privacy are not protected when derived personal information is generalised and the data mining results are used to discriminate unjustified, because it is based on an arbitrary correlation.

Summarising the above we may conclude that the current policy for protection of privacy seems inadequate due to technological developments. The focus of the WBP for protection heavily leans on protection of content, anonymising personal information and the concept of informed consent. As I will argue, these concept will all turn out to be problematic. But before we scrutinise these issues, I will first explore whether an assumption about the moral dimension of confidentiality in a physician-patient relation, patients have is still valid, because this has implications for inadequacy of the focus of the WBP to protect the content of an EHR. The focus of the WBP is namely based on a centralised database with one gatekeeper, instead of the current decentralised systems. Since we do not have one gatekeeper anymore and data is linked and available on several computers, the traditional concept of confidentiality needs to be scrutinised.

#### **4. Protection of personal data**

The moral appeals made by the public in deliberating about research participation include amongst others protection of privacy and confidentiality. Confidentiality may be seen as a subset of privacy, and most people still assume it exists the way it has been used for decades or even for centuries, while I suggest that the concept has vanished in its traditional sense and need to be replaced by another concept of confidentiality, so I will treat this concept first before I shift to protection of privacy.

##### **4.1 Confidentiality**

Confidentiality is a practice and moral duty of not disclosing information to third parties without the persons consent, when it is received in confidence. We consider information confidential when 1) we tell the



other person it is confidential and ask her to promise not to tell anyone; 2) in a close relationship we assume that information will be kept confidential, that is where a friendship and other intimate relations are based on and 3) in a professional relation, as e.g. a physician-patient relation we assume that all information we share will be kept secret. Confidentiality is allied to a direct relationship between two (or more) persons, who actually meet and see each other and exchange information. In the context of a direct relationship between physician and patient, patients still expect physicians to act on the 2500 years old Hippocratic Oath, which holds: “What I may see or hear in the course of treatment in regard to life of men, which on no account one must spread abroad, I will keep myself holding such things shameful to be spoken about”, while physicians are also bound by the international rule: “I will respect the secrets which are confided in me, even after the patient has died”<sup>15</sup> The oath and the declaration are norms physicians declare themselves bounded by, so patients still have a strong belief that information they share with their GP will never be passed on to anyone else without their explicit consent. When, in practice, one out of every twenty-five citizen, with whom there is no direct contact, have access to medical records, we can easily conclude that the traditional concept patients still have in mind has **yield** with the introduction of shared electronic health records. So, we either need to adjust our concept of confidentiality or find convincing justifications to defend the original concept of confidentiality between a GP and her patient.

A physician might justify the original norms based on a consequentialistic view that a morally right action is the action that produces the most good or utility for society. Consequentialists ask forward-looking questions. “What is the utility of this action for society in general?” is a cornerstone of their policy formation. A GP thus looks at the consequences that may occur when patients are aware that their GP can no longer live up to the norms related to the traditional bilateral confidentiality. Consequences might be that patients are unwilling to tell everything to their GP, they may hold back essential information, which makes it more difficult for a GP to make the correct diagnosis. So, it has direct negative consequences for the patient. In case of sensitive diseases, as e.g. having a sexual transmissible disease or HIV, patients may turn to private clinics and pay their own bills, so this information will not be in their medical record and this would have direct negative consequences for society as they may transfer these diseases to their sexual partners without them knowing. It might also be more costly in the end, when patients are visiting their GP too late. A physician taken a deontological approach, on the other hand, might argue that patients have to be able to trust that whatever they tell their GP, will remain confidential at all times, so advocating to make the confidentiality norm absolute. The deontological orientation is based on respect for autonomous rational persons. This has its roots in the theory of Kant, who argued that all and only persons (as rational autonomous agents) and the moral law they autonomously legislate are appropriate objects of the morally most significant attitude of respect<sup>16</sup>. To be a person is to have a status and worth that is to be an end in itself with dignity and we affirm to respect the dignity of persons as ends in themselves in our attitude and conduct. According to Kant respect for such beings is not only appropriate, but also morally and unconditionally required. Kant’s Formula of Humanity commands that our actions express due respect for the worth of persons: “Act in

---

<sup>15</sup> Declaration of Geneva, World Medical Association 1948

<sup>16</sup> <http://plato.stanford.edu/entries/respect/> (last visited 17-08-11)



such a way that you treat humanity, whether in your own person or the person of any other, never simply as a means but always at the same time as an end”<sup>17</sup>. Treating others always as an end (in themselves) is Kant’s idea of respect for autonomy. Humanity, or perhaps more precisely rational natures should be treated as intrinsically valuable. We recognise the other as having the capacity to choose her own goals and projects on the basis of moral principles known by reason. So by respecting the autonomy of others, we do not interfere with their personal conceptions of what is right. A morally right actions for e.g. the physicians are therefore those actions that express respect for persons as ends in themselves. A GP, who takes a deontological disposition based on Kant’s Formula of Humanity, might decide not to put sensitive information in the EHR of the patient. US psychiatrists<sup>18</sup> suggested to keep their records separate from the EHR, so third parties will not have access, to hold on to the traditional bilateral confidentiality concept based on respect for persons as end in themselves, but it is often of vital importance for e.g. surgeons or internists to know that a patient is seeing a psychiatrist or has some specific medication related to mental illness. If we would accept separate treatment files for mental illnesses and generalise this to other treatment session, classified by the patient or the GP as sensitive, this could lead to unacceptable consequences of having separate records for each type of medical problem and the EHR would lose its credibility because it will be incomplete.

Whether we defend the traditional direct bilateral concept of confidentiality either from a deontological perspective or a utilitarian perspective, confidentiality has *prima facie* standing. We can think of situations where we have to balance the value of confidentiality against other values. Even in the traditional bilateral physician-patient confidential relationship model the physician may find herself obliged to a moral duty to breach confidentiality and warn third persons thought of to be at risk. Examples are e.g. when a patient has a sexual transmissible disease and is unwilling to tell its sexual partners. Breaching a confidentiality norm in this case might be one way to discharge a physician of a duty of confidentiality and redirect the physician with a positive duty to protect third parties based on the moral duty of “do no harm”. In the famous Tarasoff case<sup>19</sup> (California 1976) boundaries were set to confidentiality and physicians (including psychiatrists) were addressed with a “duty to warn” those person severely at risk because of possible actions of patients under their treatment. Justice Tobriner wrote in the majority opinion in this case: “The public policy favouring protection of the confidential character of patient-psychotherapist communications must yield to the extent to which disclosure is essential to avert danger to others. The protective privilege ends where the public peril begins”<sup>20</sup>, hereby following John Stuart Mill saying that “... the only purpose for which power can be rightfully exercised over any member of a civilised community, against its will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant”<sup>21</sup>.

---

<sup>17</sup> Kant, 1993, p. 36

<sup>18</sup> [http://en.wikipedia.org/wiki/Tarasoff\\_v.\\_Regents\\_of\\_the\\_University\\_of\\_California](http://en.wikipedia.org/wiki/Tarasoff_v._Regents_of_the_University_of_California) (last visited 17-08-11)

<sup>19</sup> *ibid*

<sup>20</sup> *ibid*

<sup>21</sup> Mill, 1978, p. 43

If we hold on to a narrow concept of confidentiality between GP and patient, not passing on *any* information without explicit consent, this will have the consequence that third parties having access to someone's EHR without the patient or research participants consent is a moral failing by definition. In current health infrastructure practices the narrow concept of bilateral confidentiality has yielded with the introduction of linked computers without patients giving their explicit consent. The moral foundation to breach confidentiality in a narrow sense can be found by putting more weight on other values as e.g. beneficence or non-maleficence, doctors not being able to save people's lives in cases of emergency, because they do not have all information available straight on to make a correct diagnosis or it can be weighed against norms of justice, when third parties need to be informed of risks they encounter by having a sexual relationship with this patient without protection. Or confidentiality can be weighed against general norms of justice or solidarity, when researchers working with less or no information might jeopardise their research and will have an impact on the (lack of) improvement of health of the population. To reflect the current health infrastructure we need to adjust the traditional concept of confidentiality in a physician-patient relationship to a broader concept of multilateral indirect confidentiality, where norms to protect privacy interests and not disclose information to other parties should guide all the persons having access to the EHRs with sensitive information of the patient. Although an indirect multilateral confidentiality norm is regulated in the WBP, patients may feel weakened to place their trust and provide full information in medical settings, because they have no idea who exactly has access to their information.

An additional component is added to the problem of confidentiality nowadays. In previous practices the GP was the accountable person for the content of the EHR, due to the technical developments all kinds of persons in health care practices add information to patients EHRs. The result is that no-one is responsible (any more) for all the content, so no-one is normative accountable. But if a lot more person have access to the information in EHRs, does that mean they also have knowledge about that person? I will explain now that there is a normative difference between having information and having knowledge.

#### 4.2 Having information differs from having knowledge

The persons having access to sensitive information, which is one of every twenty-five citizen in the case of health-care, should be restricted as much as possible. Although this seems to be a lot of persons, we must take into account that those persons, who do have access to sensitive information, do not always have knowledge about us. A person knowing that a girl is 17 years old, does not have any implications, unless she also knows that this particular girl occasionally drives her mother's car, and since she can make the causal connection that 17 year old persons do not have a driver's license, she can disapprove of her action. We might think that if someone knows something about us this matter morally to us, without considering if the knowing person (the data holder) has the rational competence to make causal connections. If you do not know anything about consequences of heavily abuse of alcohol and drugs, you would not even consider this to be the cause of the early dying of Amy Winehouse<sup>22</sup>. If we intend to acquire knowledge about someone, we do so with a certain context in mind. We

---

<sup>22</sup> although we still do not have any proof that alcohol or drug abuse really is the cause of her death.

want to have knowledge for a specific reason. So, the significance knowledge has for a person, e.g. the difference it makes how one act based on that knowledge, depends upon cognitive and practical commitments. Information acquires ethical and normative significance, only because it can or may be used in certain actions. These actions always depend on background knowledge a data holder has and the competences, not just on the possession of specific information. What is informative and significant for one person might be irrelevant to others. This suggests that informational obligations bear on epistemic and communicative actions and transactions, rather than on information content. If we concentrate on insurers, who want to have access to medical information in order to increase premiums for those with higher health risks and if we see possession of information against this background, we may be led to think that information itself has some intrinsic ethical significance, which is not the case.

We may conclude that with the loss of the traditional bilateral confidentiality patients may feel weakened in their ability to maintain trust or confidence in medical or bio-medical settings. And although patients also lost one gatekeeper who is normative accountable for the content of their individual EHR, patients should be aware that persons having access to their information, do not *per se* have knowledge. We are more likely to regard certain use of information, which we morally rebuke and therefore we consider this subject to obligations. If we follow this road, we might conclude that instead of trying to establish a right to informational privacy over (sorts of) content, it would be more helpful if we could set out conditions to establish in which cases the use of information would wrong us and distinguish between morally licit and morally illicit acts of communicating and disseminating. It then seems logic to shift our focus from confidentiality to privacy interests.

### 4.3 Privacy interests

There is a broad picture of diverse types of privacy interests. Warren and Brandeis in their article “The Right to Privacy, published in 1890, argue that intrusions of privacy violate a right “to be left alone”. Other authors consider privacy as a necessary condition for human dignity, and others focus on privacy enabling intimate or interpersonal relations and also is privacy seen as a mean of controlling access to yourself<sup>23</sup>. In the case of research participants we have a data subject and a data holder who has sensitive information about the data subject. The data holder may commit privacy violations in how she processes the information and if and how she disseminates it to whom. In relation to the topic of this paper I narrow the notion of privacy down to informational privacy, which e.g. by Beauchamp and Childress is suggested as a right a person has, not to have all personal information known, used or disseminated by others without their consent and is a way of keeping control over the information about yourself, admitting that this does not cover all privacy issues. Other’s having knowledge about us morally matters for two reasons<sup>24</sup>: first, what others know about us may alter the way they act towards us. We may be treated unfairly, prejudicially, discriminated or stigmatised if someone e.g. knows our sexual orientation, religious commitment, membership of organisations, e.g. political or trade unions, medical history, potential future medical risks or socio-economic circumstances. Second, we are all subject to various

---

<sup>23</sup> <http://plato.stanford.edu/entries/privacy/> (last visited 16-08-11)

<sup>24</sup> Manson and O’Neill (2007)

emotions, like embarrassment or shame, which may be triggered if we know someone would have knowledge about specific facts on us, although this may depend on the culture we are living in, even though the attitude of the others towards us does not change. An adequate account of informational privacy should take into account these features.

#### 4.4 Personal information<sup>25</sup>

I follow Beauchamp and Childress by considering informational privacy as a right against others coming to know, publishing or disclosing personal information about another individual. Personal information can be considered as a subset of information that is true of a person but not regarded available to the general public. Classifying what exactly should be regarded as personal information is problematic. In the WBP personal information is mentioned as related to race, religion or philosophical belief, political opinion, health, criminal records, sexual preferences and membership of a trade-union, herewith following the EC Data Protection Directive<sup>26</sup>. However, what might be viewed as personal in one culture might be public knowledge in another culture. For instance, Denmark and Iceland do not have a problem with trade-union membership, this information is regarded to them as public knowledge, while France has put special obligations of confidentiality on matters as creditworthiness or debts, which is not mentioned in the EC Directive. Genetic data is part of the definition of health and sex life data in the Netherlands, Luxembourg and Portugal, while in Sweden, the processing of such data is separately regulated<sup>27</sup>. What might be personal for one person might not be classified as personal for another person. What is personal information today might be totally different in a couple of decades. Sensitive information seems to depend on the differentiation and change in prioritising of values people have, not only today, but also over time and between locations. This does not mean that a relativistic position needs to be taken and we need to classify over and over again what is sensitive for whom at which moment. We may still assume that although the kind of personal information differs in different cultures and between different persons, there is (very likely) always some information which persons want to protect, which can be classified as personal or sensitive. I think the Canadian approach is more helpful as an interpretative aid to distinguish licit from illicit actions by disclosing certain information. According to the Canadian Law sensitive data depends on two criteria: 1) the point of view of the data subject and 2) the context in which a third party uses the data<sup>28</sup>. The criteria of a data subject's point of view circumvents the problem of defining sensitive information for everyone. The context in which information is used, does address administrative situations, as e.g. raising insurance premiums. But besides that, it also covers cases of discrimination or stigmatisation based on e.g. group profiles.

#### 4.5 Group interests

Health databases are by definition collections of materials or data from groups. Sometimes the collections are disease related, other times it is a cohort of family members and it might also be ethnic related or geographical

---

<sup>25</sup> I will use personal and sensitive as interchangeable.

<sup>26</sup> Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data

<sup>27</sup> Mc Cullagh, 2006, p. 13

<sup>28</sup> Mc Cullagh, 2006, p. 16

spread. Different researchers are interested in different kind of information, and so very often the health database is set up to answer a specific question. If only part of the individual's information is available in a database, it is less interesting or even impossible to link this to another database. Researchers favour therefore data uniformity and standardisation and as much data as possible to be able to link databases for different questions. The biospecimens become even more valuable for researchers if there is a permanent link to updated EHRs, so information about new health related problems can be used directly. Not only researchers are interested however, pharmaceutical companies are also very interested. For them it is even more interesting to know which groups are at risk and for what. They are interested in a market for preventive medicine for the "future ill"<sup>29</sup>. National Health divisions are also interested in groups, since their policy is directed at groups and datamining is a helpful instrument to create group profiles.

A group profile is a set of properties or characteristics that are typically found in individuals assigned to a certain group. These properties often serve as the basis for identifying an individual with a group. Vedder (2001) detected that the kind of personal information which is contained in group profiles fall outside the sphere of protection granted to personal data. Personal data is understood as information related to an identifiable person. According to the WBP the data subject has some right with regard to her personal data. She has the right to know what data has been stored and the right to rectification. Regulating the access and possibility of rectification of the content in EHRs does not relate to the interests persons have once their data is processed and results in group profiling. Any research done by datamining does not have to be approved by any ethical commission, since there is no specific question, data is anonymised and legally available. Since group profiles are helpful in shaping policy and is a rather cheap process, it is used more and more. The results of the expansion of group profiling might be that more and more people are treated as members of a group, this might be done by public or private organisations but also by the public in general. We are used to generalise and label persons as members of a group and most of the time this does not have any particular consequences and sometimes it can have very positive consequences. But sometimes it may turn out to have negative consequences for the group as e.g. in the case of Ashkenazi Jews, who are identified as having a high probability getting Tay-Sachs disease. The distinction Vedder makes between distributional and non distributional group profiles<sup>30</sup> helps to make moral problems related to group profiles more explicit. Distributional profiles assign certain properties to a group of persons in such a way that these properties are always and unconditionally manifested by all members of that group. So, if a person has a disease called X is a member of group Y, all Y persons have the disease X. Each member of the group Y is fully aware of the fact that they have disease X. So, e.g. all grown up "little people" (group Y) are aware of the fact that they are not taller than 1.55 m. and the main reason for their short stature for most of them is that they have achondroplasia (disease X). Non-distributional group profiles are, in contrast to distributional profiles, constructed in terms of probabilities and averages. They are based on comparisons of members of the group with each other and other groups. For example, a researcher is interested in a group of persons with a

---

<sup>29</sup> <http://www.decode.me/customer-stories/bradley-bale-md-on-genetic-tests> (last visited 17-08-11)

<sup>30</sup> Vedder, 2001, p. 465

specific occupation, say shop-owners, and she is looking at the EHRs of the shop-owners by placing a data mining algorithm and discovers that these persons happen to have a more than average risk of getting a heart stroke before reaching the age of 50. The algorithm ‘discovers’ this patterns by comparing all files and search for correlations. So now we have a new group profile of shop-owners who are identified as a group most likely to get a heart stroke at an early age. The characteristics used to construct non-distributional profiles only apply to individuals in so far they are identified with this group and may have nothing to do with a particular individual. So, if a shop-owner would apply for a life insurance and is turned down merely by virtue of her association with this group, while she is in excellent condition and no heart strokes ever happened in her family, she is likely to be discriminated on the bases of a group profile. Vedder pointed that in these cases an individual is judged and treated on the bases of her belonging to the ‘wrong group of persons’<sup>31</sup>. The shop owner in contrast with a member of a distributional group profile probably has no idea that she is a member of the “wrong group”. So, the individuals who consented to provide their data (biospecimens turn into computer data) risk facing a negative impact based on (sometimes arbitrary) findings through data mining. Every person starting a shop will fall immediately in the ‘wrong group’ of persons.

Vedder uses the non-distributional profiling only for (arbitrary) constructed groups, mainly by data mining. In his analyses he leaves out group profiles of community groups, as e.g. the Ashkenazi Jews. In that case we have another division: being a member of a group you can leave or a group you cannot or do not want to leave. If I am a shop-owner, I can make a career change or sell my shop, which (hopefully) leaves me enough money to stop working altogether. At that moment I am no longer part of the ‘wrong group of persons’. But I might also remain a shop-owner, because I do not have another possibility to earn an income. Being a member of the Ashkenazi Jews, where a large part of my identity may come from I will not or cannot leave the group, in which case I will always be part of a high risk group, although as an individual I may have prove that I will never get Tay-Sachs. Suppose a researcher ‘discovers’ through datamining that women watching the television programme “House” have a high incidence of breast cancer and are refused a life insurance or they have to pay a higher premium solely on their association with some non-essential or non-causally related information, then we may conclude that this would arguably be unfair. Insurance companies who are forbidden to select persons on their actual health condition might now select on highly arbitrary information. To be able to deliberate if my higher premium is fair I need to know as an Ashkenazi Jewish shop-owner watching “House” on television, what the decisive element for the raise of my insurance premium is. We may therefore conclude that we have privacy interests viewed as individual interests on more levels, not only individual, but also as an individual member in a group. Being identified as a member of a group I have privacy interests at stake, but if we look at unfair consequences, we might turn to focus on individual interests as e.g. being treated fairly. Transparency of decisive elements in shaping policy seems to be a necessary condition when we allow that policy to be based on group profiles. However, sometimes it is not transparent and people are not aware of being a member of the ‘wrong group’. If the outcome of a research is resulting in administrative sanctions, like higher insurance premiums, this

---

<sup>31</sup> *ibid*

might turn into an unfair situation. So, it seems that constructing new non-distributional group profiles with the datamining technology can have an unfairly impact in the lives of individuals and poses serious problems for personal privacy, since their (sensitive) information is available to others if they can be “identified” as a group member. Vedder noticed that data in group profiles are often used as if they were personal data and the impact on individual persons resulting from the use of derived data can sometimes be more severe than in cases involving real personal data. I agree with Vedder, who believes that the increased production and use of group profiles may result in growing unfairness for the affected individuals. So, if we have to consider norms to regulate the use of information, these norms need to incorporate the use of information in the process of datamining which can harm individuals unfairly and disproportionate. Currently we consider our privacy to be protected by the concept of ‘informed consent’ and anonymisation, but when personal information is aggregated and subsequently used in another context to create group profiles, we may wonder if the informed consent given by research participants is still valid in a moral sense. Persons who do not want their information to be used and ‘opt out’ still might find themselves being related to a constructed group profile. If a group profile is constructed related to a postal code and if I am the only person who opt out in my neighbourhood, then there would still be a group profile, which would affect me. So, whether we provide our consent or opt out, the result would still be that we can recognise ourselves as part of ‘the wrong group’. We will therefore now turn to scrutinise the concept of ‘informed consent’.

To sum up, sensitive or personal information is hard to classify, but this does not have to be a problem if we focus on the way in which certain information used touches upon the social position and functioning of an individual. This has the advantage that we do not need to classify personal information, which might change over time and is different on different locations. In my view an intelligible approach to informational privacy and its regulation needs to argue for informational obligations that fall on those who acquire, possess, use or disclose personal data. If we can set out the demands of these informational obligations, we will have a base for arbitrating disagreements about the proper form of rights to informational privacy. The WBP legislation however is built on the view of informational privacy as a matter of rights over personal information. Its starting point is content based. Therefore the law assigns individuals’ informed consent a large role in controlling the acquisition, possession and use of their personal information. Given the breadth of the EU definition of processing (“any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction;” [art. 2 b]) it seems that anybody who examines tissue samples, or who organises or alters a genomic sequence would need to be seen as processing sensitive personal information. It seems odd that such impersonal research should be prevented or restricted given that it is not invasive nor intrusive, not prejudicial and does not affect, let alone harm any source object and given that the researcher often has no actual way of knowing who the source subject is. These problems are all consequences of trying to protect informational privacy by regulating types of informational content rather than types of actions when information is used and results are disseminated.



By steering on the moral aspects of the use of information about persons, we might have a better foundation to discuss also related problems, which do not fit with our concept of individual privacy, but do have an impact on individual's privacy. Since the WBP assigns individuals' informed consent a large role in controlling the processing of personal information, we will now turn to scrutinise this concept of 'informed consent'.

### 5. Informed consent and its inadequacy for protecting privacy in health databases

The notion 'informed consent' received detailed examination in the 1970s according to Beauchamp and Childress<sup>32</sup>. At first the notion of consenting was linked to research subjects, after the Nazi regime in World War II completely subordinated the interests of the individuals to that of science and society. Reducing risks and avoiding exploitation of individuals became the primary reason to ask for consent of a research subject. Only later on, the notion was considered to apply to medical interventions as well, thereby regulating that individuals can make an autonomous decision related to their own concept of health and can calculate the risks they are willing to venture against the benefits in a proposed treatment. While at first the focus of the concept of 'informed consent' was on researchers having a strong moral obligation to disclose all relevant information, later on the focus shifted to the extent in which patients and research subjects were able to understand the information. Expanding consent forms up to more and more pages, did not justify any consent unless a person actually clearly understood what the information was all about and which impacts possible consequences may have in its life. Because the primary justification for requirement of informed consent is to protect autonomous choice of a patient or research subject, the mean to achieve this nowadays is to request patients and research participants for their consent, which has to be informed by disclosing all information. But as I will show below, the concept of 'informed consent' is an insufficient condition for data subjects providing their biospecimens to a health database, because the information given is so general, the moral foundation of the concept, to make an autonomous decision which includes calculating risks and benefits, does not make sense.

We may start by wondering, if consenting is a sufficient justification for interventions with healthy people. Consent is recognised as justifying or legitimating acts. It removes moral objections to or liability for the performance of these acts. The maxim '*volenti non fit injuria*' (the willing person is not wronged) governs a wide range of acts<sup>33</sup>. Its binding force rests on the satisfaction of conditions of knowledge, intention, competence, voluntariness and acceptability of content. But, if consenting was ethically justified for any removal, we would have no reason to object to a market in human biospecimens<sup>34</sup>. We do not morally object to a trade market for blood, eggs and sperm and accept that there are vendors and buyers, however we do morally object to a market of organs. In third world countries some people tend to sell their kidneys to earn some money and although this is not illegal in every country, most people in western society have moral objections to this kind of market, even though these people voluntarily consented. We also morally object to persons who have a very strong need to have a completely healthy limb amputated and surgeons actually fulfilling these needs. Although these persons

---

<sup>32</sup> Beauchamp & Childress, 2001, p. 77

<sup>33</sup> The shorter routledge, p. 143

<sup>34</sup> Manson and O'Neill, 2008, p. 147



are very well informed that their limb is completely healthy, there is no medical need at all to amputate the limb, they voluntarily consent to amputate the limb. If the concept of ‘informed consent’ was a sufficient reason, then removal of healthy limbs was morally justified. So, we either accept that there is a market for organs and people who consent to remove their organs to sell is ethically permissible and justified or we view informed consent as an insufficient justification in these cases. If we agree that informed consent would be sufficient to ethically justify these actions, a person sacrificing himself just to donate all his body parts would then also be ethically permissible<sup>35</sup>. To take it one step further, it would also be ethically permissible if a person jeopardising his health by donating or selling his organs consented. And, if we accept informed consent as sufficient, we also need to accept that people can offer themselves as food, or become part of some artwork, implant silicone horns in their forehead as long as the person involved provides their consent. The informed consent would be the ethical justification for all these kind of expressions of individual’s autonomy. Since, not everyone does find this ethically acceptable, just the requirement for consent does not remove moral objections for the performance of the above mentioned acts.

The examples mentioned above clearly relate to physical interventions and are considered rather intrusive. We may analyse further whether the person’s desires are rational or irrational, rational or reasonable, first or second order desires, but this analysis will not enlighten us in examining if consenting of a healthy person in bio-medical research without intrusive interventions would be sufficient as ethical justification. We provide our biospecimens on a voluntarily basis to a health database and we can hardly call this intrusive, so our bodily integrity remains. It also may be morally justified to consent to the storage of our biospecimens and even for a long period of time, even forever, if never anything would happen with it. What remains problematic is the moral justification for the use of our biospecimens by consenting. The research community does not know for what reason it will be used, what the research question will be and in cases of datamining there is no way of predicting what patterns or correlations the data computer will discover, so what the outcome will be. Consenting in this case is quite different from the consent we give in medical settings.

Components of consent are acts of decision making and authorisation<sup>36</sup>. In medical settings, the patient approves or refuses a proposed treatment. If the patient approves, she consents to start exactly that treatment she consented to. A data subject consents to a specific research where there is a possibility to calculate risks and benefits. When biospecimens are required for a specific research question, the scope, length of the research and the duration of the storage of the tissue is known, because it is part of a well developed protocol. This requisite a truly unequivocal informed consent or refusal as is set out in the WBP. These kind of studies benefit from the concept of ‘informed consent’. This indicates that the notion “informed consent” has been introduced for very specific research plans to justify clinical trials with a physical intervention and well-defined strategies for therapeutic treatment. It became a legal justifying principle for specific research and health situations, where patients had the competence to consent, but proved over the years not to be able to cover all kinds of situations, so

---

<sup>35</sup> O’Neill, 2005, p. 149

<sup>36</sup> Beauchamp & Childress, 2001, p. 80

that now for atypical situations all kinds of derivatives of this notion have been developed<sup>37</sup>. Concepts, which are used nowadays are presumed or assumed consent, implicit or inferred, *ex post facto* consent and sufficient forms of consent as well as waivers, decisions not to decide. All these notions are constructed to keep track of all medical interventions in which case a person is not competent, unconscious, in life-threatening need or when it is too labour- or cost intensive related to the intervention etc.

The primary justification advanced for requirements of informed consent has been to protect autonomous choice<sup>38</sup>. To respect an autonomous agent means at the minimum an acknowledgement in attitude and action that a person has the right to make choices based on personal values and beliefs. This respect for autonomy is influenced by Kant, who argued that it flows from the recognition that all persons have unconditional worth, each having the capacity to determine his or her own moral destiny<sup>39</sup>. To violate a person's autonomy is treating a person merely as a mean. The demand that we treat others also as an end in themselves requires that we also foster their capacities in achieving their ends, not only avoiding treating them merely as a mean. If we allow the term 'informed consent' to remain the justificatory condition for general, undefined research, a data subject never can make an autonomous decision based on calculation of risk and benefits and the person would be merely used as a mean. If full informed consent is not possible to apply to data subjects who provide their biospecimens to a health database, we need to scrutinise if one of these derivations is and if not, we may conclude that we either need another concept or omit informed consent entirely.

## 5.1 Morally justified research without informed consent

According to the WBP bio-medical research without explicit consent is morally justifiable when information and biospecimens are 1) irreversibly anonymised and 2) lawfully held. Irreversible anonymisation creates a problem to start with. Weak forms of anonymisation do not satisfy the requirements of the WBP, while strong forms of anonymisation do not meet the needs of research. Without linked data, many lines of research, including all forms of epidemiological research and secondary data analyses will be impossible. DNA samples, even if ostensibly irreversible anonymised might still count as personal data. A researcher can distinguish each and every individual from other data subjects, even monozygotic twins, who differ at the level of single nucleotide polymorphisms (SNPs). So by possessing genetic material we possess information that *de facto* distinguishes each source subject from all other source subjects. It then follows that by possessing human biospecimens and the derived information on computers, researchers always possess identifiable information. This problem arises because WBP takes a particular exacting view of anonymisation. The legality of anonymised research using DNA samples will depend on whether or not the researcher could pick out the individual who is the source of the sample. If genomic sequencing becomes more widely used, it may be possible to form a genomic profile of the traits of the source person. This might turn out to be a variant of Bertrand Russell's theory of proper names<sup>40</sup>, where names like "Aristotle" are to be understood as implicitly descriptive; understanding a

---

<sup>37</sup> Vorstenbosch, 2007, p. 105

<sup>38</sup> Beauchamp & Childress, 2001, p. 77

<sup>39</sup> Kant, 1993, p. 44 [440]

<sup>40</sup> Called Logic atomism

proper name is a matter of knowing a list of descriptions which serve to pick out the referent. Genomic profiling would allow a *de facto* way of identifying individual subjects from their profile. So if all DNA information is potentially linkable, then none of it should be used without unequivocal consent. Without this unequivocal consent all genetic studies will be impermissible. If the WBP views any use of personal information without unequivocal consent as impermissible, even the process of anonymising will be illegal without consent of the research subject. A very recent research might clarify this even better. James Watson, one of the co-discoverers of the structure of DNA in 1953, had his whole genome sequenced and put this on the internet, available to everyone. The only thing he did not want to know himself was whether he had the ApoE-4 gene, linked to high risk of getting Alzheimer, so that information was left out. But by looking at the whole genome it was easy to distract if Watson had the ApoE gene <sup>41</sup>.

It might be objected that although in theory every person can be traced through its biospecimens, practically it is a hard job to do so. So let us assume that a researcher cannot do this without a tremendous effort in time and resources, privacy protections still remains morally problematic in cases of using legally held, anonymised data in the process of datamining. It is possible (and done) to put an algorithm in a data computer with anonymised information, which are legally held and search for correlations between e.g. colon cancer and the colour of a person's car. The outcome is a pattern that most persons with colon cancer drive a red car<sup>42</sup>. To do this search no consent is needed of any data subject and no Ethical Review Board has to approve, because the data is anonymised and there is no way what correlation will pop up. And although the WBP takes into account that administrative actions are not allowed, this is directed to individuals and not to individuals as a member of a group profile. Group profiles are made to guide policy. On the bases of group profiles health organisations and politics can and do decide to put in place guidelines or regulations, but not only they do so, also financial institutions, insurance companies and employers might use these profiles to set out their policy.

Unequivocal consent is required because researchers propose to do something what is regarded under normal circumstances as ethically or legally wrong<sup>43</sup>. WBP prohibits informational actions unless each individual subject gave their explicit consent to the use of legitimately held information about them. The consent removes the wrong associated with the practice of collecting and use of sensitive information. However, data subjects can hardly provide an unequivocal consent to let researchers use their biospecimens or EHRs, because they do not have any warrant to protect their privacy once the results are published. Data subjects can be wronged by published results, which might negatively affect them on an individual bases.

We might conclude that informed consent is not a sufficient justification for general research with biospecimens in health databases and that the rules laid down in the WBP for informational actions without consent of an individual do not even cover the process of anonymisation due to its particular exacting view of anonymisation. If consent is given to justify actions that are legally or ethically wrong without data subject's

---

<sup>41</sup> Nyholt, Yu and Visscher, 2009

<sup>42</sup> Custer, paraphrased by Tavani, 2004

<sup>43</sup> Manson and O'Neill, 2008, p. 148

consent, we still need to distinguish licit actions, which are helpful for individuals and society from illicit actions, which might harm or wrong the individual, although it may benefit society. But, as we shall see below ‘informed consent’ although not sufficient, remains important. So, what should be morally taken into account so we feel comfortable enough regarding the protection of our privacy to provide our biospecimens to health databases?

## 5.2 Additional moral safeguards

Since health care improvement is on top of the wish list of our society, we clearly favour biomedical research. It is both necessary and desirable to have more knowledge about diseases and their processes, so hopefully one day we can prevent some of them, end others or find more preferable treatments. The mean to achieve this is by doing research on large quantities of biospecimens, so it would be really helpful if more people volunteer and provide their biospecimens. However, to do so, the moral appeal made by the public for protection of privacy needs to be answered by the research community. Since the concept of ‘informed consent’ is not sufficient, although important, we have to look for additional moral safeguards to complement the moral foundation using biospecimens and derived data in health databases. We will therefore now scrutinise possible candidates as additional moral safeguards, being 1) trustworthiness of the research community, 2) additional rules for the use of information by the research community and 3) adding the concept of warrant assent and 4) conditions to the concept of ‘warrant assent’.

### 5.2.1 Trusting the research community

For rather good reasons we no longer place blind trust on the status of people, as e.g. on physicians. However, as O’Neill<sup>44</sup> points out, there are a number of cases where the Cassandra problem arises and our mistrust is misplaced. Part of the definition of trust, is that a truster must accept the risk of being betrayed, there is no guarantee. Trusting means (among other things) abandon control over the actions performed by the trustee. If there were sufficient warrants, we did not have to place trust, because in that case we were completely sure. Persons mistrusting biomedical research refer to research failures in the past and there are plenty to name.<sup>45</sup> The rigour of these failures was mainly due to the fact that healthy persons were intentionally made sick or sick person were denied medication, just to observe what happens in a natural process without interference. None of the persons consented to be part of an experiment, they were deceived or manipulated. O’Neill points out that although informed consent in cases of biomedical research is neither a sufficient nor a necessary condition, it is however important, because “it provides an important measure of protection against coercion and deception and also because it can make a distinctive contribution to the restoration of trust.”<sup>46</sup> O’Neill continues to sketch a wider context where informed consent requirements play an important role, but her context is mainly related to patients providing tissues for diagnostic treatment or specific research proposals. However, her analysis can help us as well to ground trust in the case of healthy people providing their biospecimens. If we are willing to trust the research community a procedure of ‘informed consent’ is important to avoid deception and exploitation. If we do

---

<sup>44</sup> O’Neill, 2005, p. 141

<sup>45</sup> We might think of the Tuskegee syphilis experiment, the experiment in which researchers gave hepatitis injections to healthy retarded children etc. <http://www.rbs2.com/humres.htm>

<sup>46</sup> O’Neill, 2005, p. 145

not feel deceived, we are likely to trust. For the research community it is of the utmost importance that they earn the trust by being trustworthy. One way for researchers to be generally regarded as trustworthy is to be respectful in the use of data from their data subjects, to be transparent about their research and explain to them as well as the general public their policy and procedures.

To make this more specific, researchers need to provide truthful information to potential data subjects providing their biospecimens to health databases. The truthfulness of the information to potential data subjects is often the opposite of requirements for consent in a therapeutic or clinical trial setting, which are very much focussed on a risk/benefit calculation, which needs to be proportional. The truth in case of providing biospecimens to a general health database is that information will not be kept confidential in the traditional sense, but shared by many; results will be disclosed, since publications are the aim of researchers and of benefit to society; biospecimens will not be irreversibly stripped of identifiers, unless a special request has been made; and data will be stored forever, at least as long as it is useful to the research community. Candour given information by researchers grounded in an attitude of respect for persons would add to gaining trust of volunteers.

### **5.2.2. Rules for the use of information by researchers**

One way of addressing public's moral appeal for protection by the research community is to protect data subjects against the consequences of the use of data related to them. Researchers in general have a social responsibility towards community to do research which amongst others helps to clarify patterns, find causal relations and provide insights. In biomedical research, researchers are depending on the provided information of volunteers. The more information they have the better the results, because they need to reach an appropriate sample size to make a decisive interpretation. Because they are depending on the individual volunteer and are working to benefit society, it would be consistent for them to have the interests of all individuals in society in mind<sup>47</sup>. The research practice should adopt and maintain rules that make clear that they are trustworthy and respectful, and that the public indeed trust that these attitudes are maintained. For example, such rules could include the following. Upfront, researchers may discuss in their research community, with their METC and also with the general public which kind of correlations in a datamining process might be regarded as illicit. If the research community takes the precautionary principle as moral orientation, it would imply, they take possible social impacts serious and anticipate on harm before it occurs. At the end, researchers who can imagine that their results might have an impact on individuals life, should be able to serve their draft papers to a special committee, either a new committee or (part of) the METC. This committee should at least have some lay persons on board, able to analyse if there are possible impacts for society and specific individuals and what the impacts might be, so making some kind of risks assessment. This committee should also have close relationships with the media to make sure that results which might have an impact are known to the general public. At the very least norms regulating the use of information need to incorporate the use of information in datamining processing.

---

<sup>47</sup> Although there are voices claiming that fundamental research should always be done independently and without anyone or anything in mind, this research is also bound by rules. I will not discuss this further here.

### 5.2.3. Warrant assent<sup>48</sup>

The notion “informed consent” has been introduced for very specific research plans to justify clinical trials with a physical intervention and well-defined strategies for therapeutic treatment. It became an ethical and legal justifying principle for these specific research plans and health situations, where patients had the competence to consent. We consider persons’ consent as the authorisation or permission directed to specific treatment or (intrusive) action. In these cases persons can deliberate about risks, benefits and alternatives and can make an autonomous choice. The autonomy of the person is the primary justification for asking consent. So, when a choice is made by a person and consent is provided to a specific action, we authorise the other person something to do what otherwise would be ethically or legally wrong. Consenting is done to a specific person, e.g. the GP, a surgeon or a researcher. If a patient refuses a proposed treatment, she has considered alternatives, which might differ per person, in a range of only one alternative to plenty. The choice she makes is reflecting her autonomy. Once a person gave its consent, it is directly related to that proposal.

However, in cases where we provide our biospecimens to a health database for general research, we do not do this to a specific plan or proposal. Biospecimens are used for whatever research question it might be useful for. The ethical and legal justifications in this case cannot be grounded in autonomous decision making. A data subject cannot deliberate on different options other than to join or absent oneself. We can, however, assent with to a practice of an institution. I suggest ‘warrant assent’ as a different way of thinking about how individuals can confirm their willingness to have their biospecimens stored and used for research purposes.

Opposed to informed consent, where the consent is directed to a specific action, assent implies a supporting attitude of the person to a more general practice or institution. By assenting to a biobank the data subject declares herself to be on the side of the research community and welcomes the research done in general. In analogy, persons with cancer may request to have themselves treated in the Anthony van Leeuwenhoek hospital<sup>49</sup> instead of close at home in their regional hospital or in a nearby Medical Centre. They *assent* to the way the health care providers work or think in the specific setting of that hospital and how patients are treated personally there. Perhaps we can compare this also with a membership of an organisation. Suppose I want to join a local rowing society and there are two societies in my village. And suppose, being a Christian I do not want to row on Sundays. In that case I become a member of the rowing society which plans its rowing practices on Saturdays. I assent to their practice. If they change their practices and plan all trips on Sundays, I might want to opt out. By the same analogy, a data subject assents to the way biomedical research generally takes place and welcomes the research done for the benefit of society. The data subject may expect in exchange a respectful attitude and conduct of the researchers safeguarding the interests of their data subjects. This however implies that assent is given under certain conditions - hence *warrant* assent. An explicit warrant assent given by a data subject, would address that both, data subject and data holder are aware of their responsibilities. Data subjects hand over their biospecimens, trusting the research community will also keep their personal interests in mind. Data holders

---

<sup>48</sup> Translated in Dutch: Gewaarborgde instemming

<sup>49</sup> A specific hospital only for cancer patients in Amsterdam

are aware that by the given assent, data subjects trust them and they need to act trustworthy. If the conditions for using data are made fully transparent and accountable in the form of warrants, data subjects do not have to trust the research community blindly. Explicit and transparent warrants may offer a basis for reasonable confidence that the research community will store and use their biospecimens in a respectful and responsible way.

#### 5.2.4 Conditions for warrant assent

To serve the data subject's wellbeing at best, researchers willing to publish their results might consider if these results can have a severe negative impact on an individual life, identified as member of a group. If so, they could, as a rule of conduct, send their draft publication first to their METC (Medisch Ethische Toetsing Commissie – Medical Ethical Review Board) for guidance. Members of the METC and the researcher may decide not to publish at all the way it stands now, start a public debate while publishing, approach data subjects straightforward first (if possible, otherwise through the media) or take other decisions. Researchers and members of a METC need to develop a sensitivity to search for possible negative impacts of published results. A warrant for data subjects would be to know that a researcher takes into account other persons (positive) ends in her own plans according to Kant's Kingdom of Ends<sup>50</sup> principle. If all rational beings stand under the law that they always treat themselves and others not merely as a mean but at the same time also as an end-in-themselves, then a systematic union arises of rational beings through common objective laws and this is what Kant calls a Kingdom of ends. In this kingdom of ends everything has either a price or a dignity. And "...that which constitutes the condition under which alone something can be an end in itself has not merely a relative worth, i.e. a price, but has intrinsic worth, i.e. dignity"<sup>51</sup>. Kant concludes that morality and humanity, as far as it is capable of morality, alone have dignity and therefore deserve respect. This would justify the respectful way researchers ought to operate.

If we accept the deontological constraint to use data subjects not merely as a mean, but also as an end-in-themselves, we may try to make this general imperative more specific for the research community to guide their actions when using biospecimens of data subjects for their research. If we accept that researchers have to act in such a way that the wellbeing of their data subjects are served best, they have a moral obligation to focus on the interests and concerns of a reasonable data subject and to make the assent moral valid potential data subjects should at least receive information about 1) the identity of the samples: in which kind of research is it anonymised and when not. Where is the "key" of encrypting and what are the procedures to keep samples and key separated and save; 2) the constraints attached to the form of multilateral confidentiality. Obligations attached to multilateral confidentiality in the research community merely assure that some 'direct routes' are closed off<sup>52</sup>, it does not imply that information will not become available through other routes. Furthermore, it will be a good code of conduct when infringements are sanctioned. When a pattern of moral errors of a researcher indicates a defect of moral character, we expect the research community to take this profoundly serious<sup>53</sup>. In addition, we may opt for an extra layer in monitoring the research community, because self-regulation needs to be audited.

---

<sup>50</sup> Kant, 1993, p. 39 [433]

<sup>51</sup> Ibid, p. 40 [435]

<sup>52</sup> Manson and O'Neill, 2008, p. 172

<sup>53</sup> Beauchamp & Childress, 2001, p. 31

One defect for instance, is that professionals may be too close to their fellow professionals to hold them accountable rigorously. Professional accountability has been criticised for their professional cosiness, professional capture and at worst their professional corruption<sup>54</sup>. To avoid this, we need independent persons as well as experts to be able to hold professionals accountable before results are published. Nowadays METCs are mainly at the forefront of a study, it would be worthwhile to include them as experts also at the end, before papers are published. METC members with the included lay persons should also be able to protect or at least be sensitive about current and future-oriented concerns of data subjects; 3) Data subjects should be informed about the control and ownership of the samples<sup>55</sup>; 4) They should also be aware of the limits of withdrawal from the database, because once data is used and published, there need to be a possibility for other researchers to verify or falsify the results, so they need to have access to the samples as well; 5) Of interest to a research participant is also whether there is a possibility to have personal access to relevant information derived from their specific sample, e.g. is research information added to their EHR; 56 Information should be given by researchers about the length of tissue storage. The aim of a health database is to keep this forever, or as long as it is useful; 7) Research participants should be informed that their data most likely will be used in international projects and that their data is stored on linked national and international computers, which are also used for datamining to create group profiles; 8) And last but not least data subjects direct interests are the impacts which published results may have, in cases where they are identified having a risk factor, either as an individual or as a group member and how the data subject will be protected against these risks. Since the research community cannot be held accountable for all kind of stigmatisation and discrimination cases, politics has a task in safeguarding as well.

However, data subject, as members of the general public, have a moral obligation according to Vedder to start or join the debate which correlations through data mining cannot be sought and if one agrees that correlations ending up in publication which have an effect of persons' life, are fair and or not<sup>56</sup>. Researchers and METC members should be aware of publications resulting from research and may need to be trained in sensing severe negative impact results might have on certain individuals and groups. Albeit, the community also has their own moral responsibility to keep an eye on the bio-medical publications and read annual reports of at least the health database their biospecimens are stored. And although the knowledge gap between researcher and research subject might be enormous, a genetic researcher knows an awful lot more than a volunteer to a research project, we have to keep in mind that in opposite of a physician and a patient relation, where the physician does not actually need a patient, the researcher depends on its volunteers to do research.

---

<sup>54</sup> Manson and O'Neill, 2008, p.172

<sup>55</sup> This is an entire different ethical dilemma, which will become more relevant the moment the biospecimens can be commercialized. Health databases are a research resource with significant commercial potential. It might be worthwhile for citizens or donors providing their biospecimens to request that when the health database is sold the state or university should have > 50% of the shares. I will not address this further in this paper.

<sup>56</sup> Vedder, 2000, p. 468



## 6. Conclusion

Computer technology advanced quickly during the last couple of decades. The family GP keeping track of the health status of its patients is replaced by an entire network of health care providers, all having access to a patient's electronic health record (EHRs), which includes sensitive information. Although people have different ideas about what should be regarded as sensitive data, we may assume that every person has some information which it would like to be kept secret or confidential and which has a moral worth to protect. The intention of the law on protection of personal information (WBP) intention was to protect privacy interests of the citizens in the Netherlands, but it turned out not to be watertight. Advancements in technology, as data mining, turn out not to be covered. The focus of the WBP on the person's right over content does not protect a person from disclosure of private information as a member in a created or 'discovered' data mining group profile. In this paper I have scrutinised what should be morally taken into account to protect privacy interests of research participants when they provide their biospecimens to a bio-bank, where it turns into data in a health data base. To do so, we needed to make explicit that the concept of direct bilateral confidentiality between physician and patient has yielded. Since a lot of health care providers have access nowadays to patient's EHRs, the concept should be replaced by an indirect multilateral concept of confidentiality. With the loss of confidentiality in the traditional sense, patient may feel weakened to maintain confidence in (bio)medical settings. Other's having information about us morally matters to us because 1) it may change their attitude towards us, e.g. treat us unfairly or discriminate and 2) we might feel embarrassed if we know others have information we would like to keep secret. However, patients should be aware that if someone has information that person does not have to have knowledge. If you know a person is 17 years old, then this is neutral information unless you also happen to know that she occasionally drives a car. We judge based on knowledge we have, not just on information. So, while more persons have access to information in EHRs, they do not automatically have knowledge.

I continued with examining if privacy is protected when researchers create group profiles. Health databases are collections of personal data and researchers are interested to know what these persons have in common and what divides them and this information results in group profiles. Personal information in the WBP is information related to an identifiable person, so anonymised information can be used according to the WBP. Because the WBP focuses on the content of the information, persons have a right to have access to data related to them and can ask for a rectification if they consider something not right. But they do not have anything to say about when and how the information is used. If researchers discover patterns or correlations they create group profiles. Although we are used to be labelled as e.g. house owners, you might be surprised to identify yourself as the person owing a red car having a high risk of contracting colon cancer. If your life insurance premium is higher than your neighbour owing a blue car, you probably would not even consider that the colour of your car is the decisive element if this policy is not transparent. According to the WBP privacy is protected through the concept of anonymisation. But as we saw anonymisation does not protect individuals' privacy, data subjects may always identify themselves as member of a group and so will their surroundings identify them. If a person decides to opt out, while others provide their consent, this person will still be identified as a member of the 'wrong' group. I

therefore concluded that privacy is not protected in cases of creating group profiles based on anonymised information.

Informed consent has been introduced to ethically and legally justify invasive actions in medical treatment and research, e.g. clinical trials. These actions have a clear scope and aim and the concept is directed towards competent adults. The primary reason to request consent was based on the patients' right to make autonomous decisions. Over the years this concept of informed consent has proved not to cover atypical situations, so all kinds of derivatives have been introduced. When biospecimens are anonymised and legally held no consent is required according to the WBP, while in cases of data mining results can have a severe moral impact on people's life.

So relying alone on the informed consent concept does not seem sufficient. We may conclude that additional moral safeguards have to be put in place to protect privacy interests if persons provide their biospecimens for general research. I therefore propose to use the notion of 'warrant assent' as a different way of thinking how individuals can confirm their willingness to have their biospecimens used for general research questions.

"Assenting" implies an explicitly supporting attitude to a more general practice or institution instead of consent to a specific research proposal. A data subject may assent to a biobank and agree with the storage and use of her biospecimens, and thus confirm that she accepts its way of working. The data subject may expect in exchange a respectful attitude and conduct of the researchers safeguarding the interests of their data subjects. This however implies that assent is given under certain conditions - hence *warrant* assent.

To serve the data subject's wellbeing at best, researchers willing to publish their results might consider if these results can have a severe negative impact on an individual life, identified as member of a group. If so, they could, as a rule of conduct, send their draft publication first to their METC (Medisch Ethische Toetsing Commissie – Medical Ethical Review Board) for guidance. Members of the METC and the researcher may decide not to publish at all the way it stands now, start a public debate while publishing, approach data subjects straightforward first (if possible, otherwise through the media) or take other decisions. Researchers and members of a METC need to develop a sensitivity to search for possible negative impacts of published results.

If we accept that researchers have to act in such a way that the wellbeing of their data subjects are served best, they have a moral obligation to focus on the interests and concerns of a reasonable data subject. Data subjects might be willing to assent if warrants are put in place. I suggest eight warrants as requirements to validate our assent so researchers are ethically justified to use our biospecimens for general research. These requirements are 1) a notification that biospecimens are not irreversible anonymised; 2) the constraints belonging to the concept of multilateral confidentiality; 3) rules applying to control and ownership of data; 4) awareness of the limits of withdrawal; 5) clear information about the length of storage, which might be forever; 6) if and how persons have an opportunity to learn about information derived from their samples; 7) information about

international cooperation between research institutes and their sharing of data and finally 8) if and how are research participants protected against risks of belonging to the ‘wrong’ group in cases of group profiling. Research participants should be aware that group profiles are made to shape policy and are most of the time very helpful, but sometimes being an identified member in a group profile can be very harmful for the individual. Development of technology will continue, so it is also up to society to distinguish licit from illicit actions in research. Society needs to interfere when they regard policies to be unfair. Based on values commonly shared in society, a conclusion may be drawn that the research community cannot ask certain questions. As Hans Jonas cautioned: “Let us also remember that a slower progress in the conquest of disease would not threaten society, grievous as it is to those who have to deplore that their particular disease be not yet conquered, but that society would indeed be threatened by the erosion of those moral values whose loss, possible caused by too ruthless a pursuit of scientific progress, would make its most dazzling triumphs not worth having.”<sup>57</sup> We consider biomedical research highly important and by putting warrants in place we may trust the research community enough to assent and provide our biospecimens for general research for the benefit of society. But society also has a moral obligation to audit the research community.

***Doveryai, no proveryai!***

Russian proverb  
(*Trust, but verify*)

---

<sup>57</sup> Jonas, 1969, p. 246

## Literature

Annas GJ et al. Drafting the genetic privacy act: Science, policy and practical considerations, *Journal of Law, Medicine and Ethics* 23: 360-6, 1995

Arnason E, Personal identifiability in the Icelandic health sector database, *Electronic Law Journals – JILT* 2: (last visited 15-04-11)

Ashcroft R, The ethics of reusing archived tissue for research, *Neuropathology and Applied Neurobiology* 26: 408-11, 2000

Bedau MA and Parke EC (eds), *The Ethics of Protocells, Moral and Social Implications of Creating Life in the Laboratory*, The Precautionary Principle and Its Critics, MIT Press, Cambridge Massachusetts, London, UK 2009

Beauchamp TL and Childress JF, *Principles of Biomedical Ethics*, fifth edition, Oxford University Press, Oxford, 2001

Becker M, Stokkom B van, Tongeren P van, Wils JP (eds) *Lexicon van de ethiek*, Koninklijke Van Gorcum BV, Assen. 2007, Notion “Informed Consent” by Vorstenbosch J. p. 165-166

Burley J and Harris J (ed), *A Companion to Genethics*, Blackwell Publishers Ltd, USA, Powers M, Privacy and Genetics, 2002

Caplan AL, Moreno JD, The Havasu ‘Baaja tribe and informed consent, [www.thelancet.com](http://www.thelancet.com) 377:621-622, 2011

Craig E (ed), *The shorter routledge encyclopedia of philosophy*, Routledge London and New York, 2005

Fule P and Roddick J, Detecting Privacy and Ethical sensitivity in Data Mining Results, Australian Computer Society Inc, *Conferences in Research and Practices in Information Technology*, Vol. 26 Estivill-Castro V (ed), 2004

Frankena WK, *Fundamentele Ethiek*, translated and published by Van Gorkum & Comp. BV, Assen, 1977

Jonas H, Philosophical reflections on experimenting with human subjects. *Daedalus* : 98, 219-247, 1969

Kant I, *Grounding for the Metaphysics of Moral, On a supposed Right to Lie because of Philanthropic Concerns* (1785), translated by Ellington JW, Hackett Publishing Company Inc, Indianapolis/Cambridge, third edition, 1993

Lunshof JE, Chadwick R, Vorhaus D and Church G, From Genetic Privacy to open consent, *Nature Reviews Genetics*, AOP, published online 1 April 2008 (doi:10.138/nrg2360)

Manson NC and O’Neill O, *Rethinking Informed Consent in Bioethics*, Cambridge University Press, Cambridge, reprinted 2008

Mc Cullagh K, *A study of data protection, harmonization or confusion?* 21<sup>st</sup> Bileta Conference: Globalisation and Harmonisation in Technology Law, Malta, April 2006

Mill JS, *Over Vrijheid*, translation by Krul WE, Boom Meppel, Amsterdam, 1978

Murray T, ‘Genetic Exceptionalism and “Future Diaries”’: Is Genetic Information Different From Other Medical Information?’ in M Rothstein (ed), *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era* (1997) Yale University Press, New Haven, 1997

- Nyholt DR., Yu CE and Visscher PM, On Jim Watson's APOE status: genetic information is hard to hide, *European Journal of Human Genetics* Macmillan Publishers Limited 17, 147–150 2009
- O'Neill O, *A question of trust*, The BBC Reith Lectures 2002 Cambridge University Press, Cambridge, reprinted 2003
- O'Neill O, *Autonomy and Trust in Bioethics*, Cambridge University Press, Cambridge, third printing 2005
- Pira E et al, Cancer mortality in a cohort of asbestos textile workers, *British Journal of Cancer* 92: 580-6, 2005
- Ploem M.C., *Tussen privacy en wetenschapsvrijheid. Regulering van gegevensverwerking voor medisch-wetenschappelijk onderzoek*, Dissertatie 2004: <http://dare.uva.nl/record/205805>
- Poschl G and Seitz HK, Alcohol and cancer. *Alcohol and Alcoholism* 39: 155-165, 2004
- Roche PA and Annas GJ, Protecting genetic privacy. [www.nature.com/reviews/genetics](http://www.nature.com/reviews/genetics) *MacMillan Magazine Ltd.* volume 2: 392-6, May 2001
- Sankar P, Genetic Privacy, *Annu Rev. Med.* 54: 393-407, 2003
- Schneewind JB, *The invention of autonomy: A history of modern philosophy*. Cambridge University Press, New York, 1998
- Skrikerud AM, The Dubious Uniqueness of Genetic Information, in Solbakk et al (eds), *The Ethics of Research Biobanking*, 57-67, 2009 (DOI 10.1007/978-0-387-93872-1.5)
- Solbakk JH et al. (eds), *The Ethics of Research Biobanking*, Skrikerud AM, The Dubious Uniqueness of Genetic Information, p. 57-67, Springer Science + Business Media LLC, 2009
- Spaans JAJ, Hoonakker DGPH, Leer van der OFC (red), *Privacy, een kind van vele ouders*, Ned. Ver. Van Medische Audiovisuele Communicatie, (NVMAC) Oss, 1997
- Tavani HT, Genomic research and data-mining technology: Implications for personal privacy and informed consent, *Ethics and Information Technology* 6: 15-28, 2004
- Tschantz MC, Wing JM, Formal Methods for Privacy, Cavalcanti A and Dams D (eds), Springer-Verlag Berlin Heidelberg, *FM 2009 LNCS 5850*, p.1-15, 2009
- Ursin LØ, Duties and Rights of Biobank Participants: Principled Autonomy, Consent, Voluntariness and Privacy, in Solbakk JH et al (eds.) *The Ethics of Research Biobanking (Springer Science + Business Media LLC)*, 2009 (DOI 10.1007/978-0-387-93872-1\_6)
- Vedder A, Medical Data, New Information Technologies and the Need for Normative Principles Other than Privacy Rules, In M. Freeman & A. Lewis (eds.), *Law and Medicine : Current Legal Issues: Volume 3* (pp. 441-459). Oxford: Oxford University Press. (Current Legal Issues, 3), 2000
- Vedder AH, Het einde van de individualiteit? Datamining, groepsprofilering en de vermeerdering van brute pech en dom geluk. *P&I Privacy en Informatie*, 1; 3: 115-120, 1998
- Vineis P et al, Misconceptions about the use of genetic tests in populations, *Lancet* 357: 709-12, 2001
- Zwenne GJ, Duthler AW, Groothuis M, Kielman H, Koelewijn W and Mommers L, Eerste fase evaluatie Wet bescherming persoonsgegevens, *Literatuuronderzoek en knelpuntenanalyse*, 2007

**Consulted websites:**

<http://plato.stanford.edu/entries/privacy/>

<http://www.decodeme.com/customer-stories/bradley-bale-md-on-genetic-tests>

<http://www.youtube.com/watch?v=sf0YXnAFBs8>

[http://www.ornl.gov/sci/techresources/Human\\_Genome/home.shtml](http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml)

[http://www.ornl.gov/sci/techresources/Human\\_Genome/elsi/elsi.shtml](http://www.ornl.gov/sci/techresources/Human_Genome/elsi/elsi.shtml)

<http://plato.stanford.edu/entries/respect/>

<http://www.rbs2.com/humres.htm>

With special thanks to Prof. Peter Heutink, Head of the department Medical Genomics at VUmc, Amsterdam, for his insight information on the processes of genetic research and background information about genetics.