Prof. Bernice Elger (University of Geneva)

Ethical and Legal issues concerning sharing of medical information across European borders

Prague, May 15, 2009

Overview of the presentation

Introduction

Why we need to exchange medical information online?
Cases: KnowARC and @neurIST

Ethical and legal issues

Guidelines
Privacy and degrees of anonymity
Informed consent - impossibility?

Hot topics Unresolved issues, working groups

Need for information exchange

- o Research on rare diseases
 - Even a major hospital may see only a case or two
 - Sufficient expertise does not exist in any single place
- o Epidemiological studies
 - Contagiousness, spread, mortality initially impossible to determine locally
 - Case: A H1N1
- o Mobility
 - Patients and doctors move across national borders

tegrated Biomedical Informatics for the Management of Cerebral Aneurysms - Windows Internet Explorer

http://www.cilab.upf.edu/aneurist1/	💙 🗲 🗙 Google
Affichage Favoris Outils ?	
neurIST - Integrated Biomedical Informa	🖓 🔹 🔝 🔹 🖶 Page 🔻

uhaite exécuter le module complémentaire « QuickTime » publié par « Apple Computer, Inc. ». Cliquez ici si vous faites confiance à ce site et à ce module pour l'autoriser à s'ex



tegrated Biomedical Informatics for the Management of Cerebral Aneurysms - Windows Internet Explorer

🥖 🕑 📁 👿 💽

http://www.cilab.upf.edu/aneurist1/index.php?option=com_content&task=view&id=38&Iter	iid=57 🖌 🖌 Google
Affichage Favoris Outils ?	
neurIST - Integrated Biomedical Informa	🐴 • 🔊 • 🖶 • 🔂 Page 🕶
uhaite exécuter le module complémentaire « QuickTime » publié par « Apple Computer, Inc. »	Cliquez ici si vous faites confiance à ce site et à ce module pour l'autoriser à s'ex

d Integrated biomedical informatics for the management of cerebral aneurysms LOGIN ? Username Password Objectives HOME LATEST NEWS GENERAL OVERVIEW Summary 13-14/3 Annual Objectives Review @UPF in Benefits Barcelona PROJECT ACTIVITIES @neurIST will provide an IT infrastructure for the management, integration and processing @neuLink of data associated with the diagnosis and treatment of cerebral aneurysm and subarachnoid haemorrage. This infrastructure will: @neuFuse @neuRisk · Facilitate clinicians the diagnosis and study of the disease, as a result of @neuEndo providing a seamless access to patient data using data fusion and processing @neuInfo of complex information spanning from the molecular to the personal level. @neuCompute · Provide a better planning and personalisation of minimally invasive interventional PARTNERS procedures for patients, after linking modern diagnostic imaging to computational tools. **NEWS & DOCUMENTS** · Collaborate in the development, extension and exploitation of standards and Newsroom protocols at all project stages. Newsletter · Share biomedical knowledge providing access to a set of software tools and Press Releases platforms such as @neuLink, @neuFuse, @neuRisk, @neuEndo, Press Package @neuCompute and @neuInfo. Publications · Create awareness through scientific dissemination and collaboration. CONTACT US · Explore the business opportunities directly arising from @neurIST. PATIENTS CLINICIANS 🙆 Internet 🖉 3 Internet Ex... 🔻 🚞 C:\HUG 9.6.04\... 👜 aneurist - Micros... 🧕 2 Microsoft Of... 🔻 aaa 2006 🎽 aaa 2006 🎽 @neurIST is focussed on cerebral aneurysms and intends to provide an integrated decision support system to assess the risk of aneurysm rupture in patients and to optimize their treatments.

@neurIST believes that the current process of cerebral aneurysm diagnosis, treatment planning and treatment development is highly compromised by the fragmentation of relevant data.

@neurIST presents a new paradigm to understand and manage cerebral aneurysms. A complete IT infrastructure will be developed for the management and processing of the vast amount of heterogeneous data acquired during diagnosis.

@neurIST also adds genetic data (from blood and tissue)



Cases: KnowARC



Risk mitigation



Support from multiple sources *NDGF, NorduGrid, KnowARC,...*Multi-discipline production use since 2001
Light-weight, non-intrusive, standards- and component-based Grid middleware *Maximise future platform choises*



MedTing – arcGIFT integration

- MedTing social web platform for medical image sharing
- GIFT GNU Image Finding for content-based image retrieval (visual similarity-based)
- Integration using HED (Hosting Environment Daemon) as web service interface



Due diligence supporting ARC

- Requirements
 - More advanced functionality than just cycle scavenging
 - Legal and regulatory constraints
 - No uncertified changes to underlying IT infrastructure -> need to run on "anything"
 - "Zero additional load" for IT support
 - Developer automates, users monitor
- Interoperability
 - Allow integrating heterogeneous resources
- Knowledge available in-house
 - KnowARC project

Cases: KnowARC @ Unige

- Research problem
 - What earlier case does the patient image (x-ray, MRI, CT, photograph) resemble?
- o Constraints
 - Patient information can't leave the hospital
 - Underlying heterogeneous IT-infrastructure cannot be changed and IT department can't accept new responsibilities
- o Solution
 - Virtualised and automated ARC installation analysing images on case reports (Internet, collaborating hospitals)
 - Search index stored only in authorised organisations' premises

Common themes in both cases

• New and partly unclear legal issues in addition to technology development

Strategies to minimize technology risk

 – e.g. minimize risk to privacy through PET, privacy enhancing technology

 at the same time: facilitating exchange of information is for the benefit of medical science and helps patients

Acknowledgements

Projects making this presentation possible
 – KnowARC – e-IRGSP2

Ethical and legal questions

Ethical and legal issues

- Laws and guidelines
- Informed consent to data use
- Privacy and degrees of anonymity (personal or not identifiable data)
- Hot topics (a global/European data protection, ongoing debate: working groups...)

Guidelines...

National databases

CDC NHANES III

Countries and health providers are following Iceland's path and combining health and genetic data on large populations. They promise to deliver "personalized" medicine, but will they?

Population Databases <u>Boom</u>, From Iceland to the U.S.

SOME PROPOSED POPULATION DATABASES

Project	Company	DNA Sample Size	Budget	Status
Icelandic Health Sector Database	deCODE Genetics	280,000	\$212M	Health database in 2003?; 80,000 DNA samples genotyped
Estonian Genome Project	EGeen Internationa	al 1,000,000	\$150M	3-year, \$2.5M pilot (10,000 donors) began fall 2002
BioBank UK	?	500,000	\$66M	Full enrollment in 2004
Marshfield Personalized Medicine	-	40,000	\$3.8M+	Enrolling this fall
National Children's Study	-	100,000	?	Full study begins in 2004
Latvian Genome Database	?	60,000 pilot	\$1.7M	Law passed in June; seeking funding
Quebec CARTaGENE	?	50,000+	\$19M	Seeking funding
	EXISTIN	G BIOBANKS AND/	OR HEALTH	RECORDS
Västerbotten, Sweden	UmanGenomics	80,000	-	Data use agreement with county in 2002
Mayo Clinic	?	100,000	-	Prototype health database completed in July
EPIC	-	350,000	-	Pooling data for cancer studies through consortium
Nurses' Health Study	-	63,000	-	
American Cancer Society CPS-II	-	110,000	-	0

7300

1160

_

Proposals to use individual data requested fall 2002

8 NOVEMBER 2002 VOL 298 SCIENCE

Regulation genetic databases

Laws

Position papers of national committees

Supranational recommendations

Genetic data: UNESCO 2003 Genetic databases:WHO 2001 DNA sampling: HUGO 1998 Biolog. material: COE 2006.

Iceland, Estonia, Sweden

Biobanks: Canada 2003,

France 2003, Germany 2004

Human biological material:

USA 1999, UK MRC 2001

Professional organisations

DNA sampling: ASHG, EuSHG Tissue: Pathologists (US, UK)

Data protection laws/recommendations

Europe

- RECOMMENDATION No. R (97) OF THE COMMITTEE OF MINISTERS TO MEMBER STATES ON THE PROTECTION OF MEDICAL DATA
- DIRECTIVE 95/46/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 24 October 1995

on the protection of individuals with regard to the processing of personal data and on the free movement of such data

Data protection laws of different countries

- Problem: are data protection laws of different countries compatible?
- In general: when data are transferred to another country, the data protection must be at least of the same level as in the country where the data have been collected.
- @neurIST: used the highest standards, i.e. from the country that had the strictest legal requirements

Definition "biobank" (genetic database)

Biobank: Tissue AND data Should be distinguished:

- The storage of the tissue samples
- The storage of the data linked to the samples/obtained from the samples
 - Information concerning the donor of the material: demographic characteristics, type of the disease, outcome of the disease, treatment etc.
 - DNA = information → Genetic database



Consent to the collection and use of data

Rules:

- If research/use of information is carried out with identifiable information, the informed consent of the patient/person is needed.
- If research/use of information is carried out with nonidentifiable information, in principle a right to "opt out" (say "no") should exist.
- Problem of identifiablity through cross-linking of different databases.

Consent to the collection and use of data

Risks:

- Patients must be informed about the risks they take when agreeing to their information being transferred and shared across countries.
- How can risks be defined? According to risks of e-banking etc.?

• How secure are the protections (PET)?

Database use and informed consent

A major problem of prospective databases is related to the issue of informed consent:

- The doctrine of informed consent is fundamental in classical health research ethics (Nuremberg, Helsinki etc.).
- Consent to "future research projects on data" is not informed, because the details of the future project are not known.

Informed consent, a burden ?

If narrow informed consent is maintained...

Anonymisation:

A solution to escape the burden of informed consent ?

Anonymisation

o Declaration of Helsinki:

Research on non-identifiable human material or nonidentifiable data is NOT medical research involving human subjects.

- No obligation to obtain informed consent.
- No obligation to obtain approval of the protocol from a research ethics committee (REC).

Biobanks: anonymisation

Non-identifiable =



Anonymisation and confidentiality...

Genetic data and confidentiality

Icelandic database shelved as court judges privacy in peril

NATURE|VOL 429|13 MAY 2004|www.nature.com/nature

Iceland's supreme court has ruled that the transfer of a dead patient's health data to a proposed genetic database would infringe the privacy rights of the man's descendants. The ruling has been interpreted to mean that the 1998 law governing the creation of the database is unconstitutional because it fails to protect personal privacy adequately.

The database plan was controversial in Iceland: more than 20,000 people actively opted out of it.

Genetic data and confidentiality

Genetic Privacy

Japan Guidelines Under Fire After Protest Halts Study



www.sciencemag.org SCIENCE VOL 301 22 AUGUST 2003

Confidentiality

Why is confidentiality important?

- The risks of a research/banking project in which
 DATA are used are for the person (whose data are used) principally related to a lack of confidentiality:
 - Discrimination (loss of insurance, employment etc.)
 - Stigmatisation
- Protection of confidentiality
 - Laws violation of confidentiality is punishable (Art. 321 Swiss CP)
 - Laws genetic discrimination is forbidden (Swiss Law LAGH)
 - Anonymisation

« Anonymisation »

 Terms related to different degrees of « anonymisation » in various guidelines relevant to research involving data vary widely.

Terms used (data or samples)

Anonymous Anonymised Anonymously coded Unidentified **De-identified De-linked Permanently de-linked** Irreversiblement anonymisé Not traceable Irretrievably unlinked to an identifiable person (UNESCO) Completely anonymised Unlinked anonymised Traceable Réversiblement anonymisé Coded

Identifiably linked **Pseudonomised** Unlinked Unlinked to an identifiable person (UNESCO) Encoded Encrypted Identified (NBAC) Nominative Directly identified (Clayton et al 1995) Fully identifiable Confidential (NHS Confidentiality Strategy) Linked to an identifiable person (UNESCO) Identifiable Personal data

Terminology: « anonymisation »

• A tower of Babel:

- A multitude of different terms. Almost for each guideline, a separate terminology is used (although there are some "traditions": UK terminology has been used in the European guidelines)
- The same term is used with a different meaning: Caution: "anonymised" and "coded" !

Terminology: "anonymisation"

European guidelines (Council of Europe)

- Anonymous
- Unlinked anonymised (Irréversiblement anonymisé)
- Linked anonymised (Réversiblement anonymisé)
- Coded
- Identified (name, address...)

Terminology: "anonymisation"

• Anonymous: data that cannot be linked to a person (difference between anonymous and anonymised not always made)

Terminology: « anonymisation »

- Anonymised: patient record (e.g. name, address, type of tumor, received treatment, donor's age etc.), but all information that would permit the identification of the donor/patient is stripped...
 - Irreversibly (unlinked)
 - Reversibly (linked): identification is possible via a code (pseudonym), but researchers/users of the data don't have access to the code

- Coded: = linked (reversibly) anonymised, but difference: researchers/users have access to the code.
- Identified: the information that permits the identification (name, address etc.) is directly associated with the dataset.

Terminology: « anonymisation »





Council of Europe 2006

Non identifiable

• Unlinked anonymised (irréversiblement anonymisé)

• Linked anonymised (réversiblement anonymisé)

Coded

Identifiable

Identifiable (personal data)

Identifiable (personal) data require higher degrees of protection and consent for use

Controversy: non identifiable data: in the case of medical data different possibilities (opt out, should patients know that their data will be used in anonymous form?)

Confidentiality and anonymisation: most important controversies

a. In which form should data and samples

- be stored ?
- be used by researchers ?
- b. Who decides which degree of anonymisation is adequate/sufficient?
- c. How many characteristics must be stripped to obtain true irreversible or reversible anonymisation?
- d. Standards for technical questions of security: firewalls, data in computers without access to any network ?

Confidentiality and anonymisation: The most important controversies

Controversy: How to assure anonymisation (linked or unlinked) that is appropriate/sufficient to guarantee that a person/patient is not identifiable?

Confidentiality and anonymisation

A person can be

- Directly identifiable (name, address, social security number)
- Indirectly identifiable, e.g. through a combination of rare characteristics (rare tumour, rare profession: according to the size of the sample population, two or three characteristics could define a person: "the university professor in Bern suffering from glioblastoma).

Confidentiality and anonymisation

Medical privacy (HIPAA) legislation USA:

- Is the method for obtaining anonymisation efficient/sufficient?

Rule: A person with appropriate knowledge and experience in statistics and scientific methodology determines that the risk of identification is "minimal" and provides documentation on the methods how he/she arrived at this conclusion. (Kulynych and Korn N Engl J Med 2002 Oct 10 p. 1133-1134).

Personal data

 RECOMMENDATION No. R (97) OF THE COMMITTEE OF MINISTERS TO MEMBER STATES ON THE PROTECTION OF MEDICAL DATA

> "- the expression 'personal data' covers any information relating to an identified or identifiable individual. An individual shall not be regarded as "identifiable" if identification requires an unreasonable amount of time and manpower. In cases where the individual is not identifiable, the data are referred to as anonymous;"

Personal data

- DIRECTIVE 95/46/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data
- NOT "unreasonable amount of time and manpower"
- "'personal data' shall mean any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity";

@neurlST

What is Anonymisation and Pseudonymisation?

Anonymisation and Pseudonymisation are vital concepts to protect both a user's and a subject's privacy. They fall in the category of so-called Privacy Enhancing Technologies (PET) and are applied in various application domains including for example communication and e-health. Unfortunately, a lot of different definitions of these terms exist including those describing different degrees of anonymisation or pseudonymisation...

@neurlST

What is Anonymisation and Pseudonymisation?

...Different definitions of the terms anonymisation and pseudonymisation lead to ambiguous interpretations and makes it difficult to have a common understanding.

That's why we in @neurIST developed our own illustration in accordance with the content of the mentioned standards and initiatives...



Anonymisation of x-ray images

Hot topics: - definition of adequate anonymisation (EC working group)

WP 136: European Commission 20 Juin 2007

Anonymous data

"Anonymous data" in the sense of the Directive can be defined as any information relating to a natural person where the person cannot be identified, whether by the data controller or by any other person, *taking account of all the means likely reasonably to be used either by the controller or by any other person* to identify that individual."

WP 136: Commission européenne. 20 Juin 2007

"Anonymised data" would therefore be anonymous data that previously referred to an identifiable person, but where that identification is no longer possible. [...]

Again, the assessment of whether the data allow identification of an individual, and whether the information can be considered as anonymous or not depends on the circumstances, and a case-by-case analysis should be carried out with particular reference to the extent that the means are likely reasonably to be used for identification [...].

This is particularly relevant in the case of statistical information, where despite the fact that the information may be presented as aggregated data, the original sample is not sufficiently large and other pieces of information may enable the identification of individuals."

http://ec.europa.eu/justice_home/fsj/privacy/workinggroup/wpdocs/2007_en.htm

WP 136: Commission européenne. 20 Juin 2007

"In other areas of research or of the same project, reidentification of the data subject may have been excluded in the design of protocols and procedure, for instance because there is no therapeutical aspects involved.

For technical or other reasons, there may still be a way to find out to what persons correspond what clinical data, but the identification is not supposed or expected to take place under any circumstance, and appropriate technical measures (e.g. cryptographic, irreversible hashing) have been put in place to prevent that from happening...

WP 136: Commission européenne. 20 Juin 2007

...In this case, even if identification of certain data subjects may take place despite all those protocols and measures (due to unforeseeable circumstances such as accidental matching of qualities of the data subject that reveal his/her identity), the information processed by the original controller may not be considered to relate to identified or identifiable individuals taking account of *all the means likely reasonably to be used by the controller or by any other person.*"

WP 136: Commission européenne. 20 Juin 2007

Pseudonomisation in the **DATA PROTECTION WORKING PARTY** text is used in a similar way as "coded", not necessarily strictly "reversibly anonymised" (where users/researchers do not have access to the code).

"Pseudonymisation can be done in a retraceable way by using correspondence lists for identities and their pseudonyms or by using two-way cryptography alorythms for pseudonymisation. Disguising identities can also be done in a way that no reidentification is possible, e.g. by one-way cryptography, which creates in general anonymised data."

Confidentiality and anonymisation

d. Technical questions of security: passwords, firewalls, data in computers without access to any network.

- Standards need to be defined

 Informatics department of institutions should be contacted when establishing a bank (used for research) to assure maximum local standards available.

WP 136: Commission européenne. 20 Juin 2007

- "The effectiveness of the pseudonymisation procedure depends on a number of factors (at which stage it is used, how secure it is against reverse tracing, the size of the population in which the individual is concealed, the ability to link individual transactions or records to the same person, etc.).
- Pseudonyms should be random and unpredictable. The number of pseudonyms possible should be so large that the same pseudonym is never randomly selected twice. If a high level of security is required, the set of potential pseudonyms must be at least equal to the range of values of secure cryptographic hash functions."

Genetic databases: anonymisation

- How many codes ?
- Independent third party ?
- Is simple coding sufficient ?

Double coding and independent third party



Confidentiality/anonymisation

Facilitate access to genetic databanks for researchers:

Genetic databanks should be put in the « public domain » (Internet etc.)

Specifying DNA sequence at only 30 to 80 statistically independent **SNP** positions will uniquely define a single person

Unrelated persons differ in about 0.1% of the 3.2 billion bases in their genomes (3). Now, the most widely used forms of forensic identification rely on only 13 to 15 locations on the genome with variable repeats (4, 5). Single nucleotide polymorphisms (SNPs) contain information that can be used to identify individuals (5, 6). If someone has access to individual genetic data and performs matches to public SNP data, a small set of SNPs could lead to successful matching and identification of the individual. In such a case, the rest of the genotypic, phenotypic, and other information linked to that individual in public records would also become available.

Conclusions

General conclusions

What would be necessary for a European research infrastructure including medical data?

- When developing of software/middleware and other technological means, one should bare in mind the growing importance of transfer/use etc. of medical and other sensitive information.
- Systems should be made compatible and suitable for the use/transfer etc. of medical or other sensitive information

General conclusions

- It is evident that the lack of consensus concerning anonymisation terminology and technology interferes with the efficiency of research.
- In order to maximize the benefit of genetic and other medical databases, it is important to put in place a well defined framework of information transfer and PET technology.
- The framework./ technology should be as global as possible in order to facilitate international collaboration.
- The risks of identification must be explored and defined.

General conclusions

- Technologies must be described qualitatively and quantitatively in terms of the degree of data protection.
- Make sure adequate patient consent has been obtained for projects involving data.
- In countries with relatively low legal standards and requirements, informed consent and high level protection should still be used in order to ensure international compatibility and the possibility to exchange data and to be "trustworthy" for other countries with stricter requirements.

The Ethics and Regulation of Human Genetic Databases

Global Perspectives

Edited by Bernice Elger, Nikola Biller-Andorno, Alexandre Mauron and Alexander M. Capron



Book
publication:
Ethical Issues
in Governing
Biobanks:
Global
Perspectives,
Ashgate
(autumn 2008)