Transparency & Access to documents: what it means to patients

François Houÿez
Director of Treatment Information & Access @ Eurordis

ADEDIEM 14th Conference, Châtenay-Malabry, 3/02/2014
CT volunteers in Europe: what can we do to increase trust in research?

• Grant open access to data to all?
• Lower data privacy standards?
• Lower the value of the informed consent?

• From EMA report comparing 2005 and 2011(1):
  – % of volunteers recruited in EU / EEA / EFTA decreased
    • from 37.0% to 31.2%
  – The contribution of the 15 MS that were already MS before May 2004 plus Norway, Iceland and Liechtenstein dropped
    • from 32.1% to 19.4%
  – The contribution from countries that became EU MS in 2004 and 2007
    • Increased from 4.7% to 11.7%

(1) Clinical trials submitted in marketing authorisation applications to the European Medicines Agency
Overview of patient recruitment and the geographical location of investigator sites
Data protection often an obstacle for sharing data for research purposes. Full respect of consent essential for public trust in research. Data protection should not be lifted just to relieve researchers from technical constraints.
Advantages of Sharing Data for the R&D

• **Efficacy data would be better due to better...**
  – mixed treatment comparisons when head-to-head data not available
  – critical mass when pooling the data for data mining activities
  – health economic models
  – massive studies examining the long-term effects of different drugs and treatment options

• **Public health knowledge increases due to better..**
  – knowledge of the disease
  – knowledge of the patient-driven effect
  – knowledge of the public health impact of different environmental factors or different lifestyle choices
Risks of Sharing Data

• Accessing data = identifying individuals becomes easier
  – Some are more vulnerable than others
  – Medical data in general are highly “delicate”
  – France « La loi Informatique et Libertés encadre la collecte et le traitement de ces données sensibles, car leur divulgation ou leur mauvaise utilisation peut porter atteinte à l’intimité ou à la vie privée des personnes concernées. Elle assure une protection renforcée de ces informations. » (Cnil.fr)
    • Les données relatives à la santé sont considérées par la loi Informatique et Libertés (article 8) comme des données sensibles dont le traitement et la collecte sont par principe interdits.
    • Des dérogations à ce principe existent
Which type of envelope do many patients prefer for their support group magazine?
Cult to transparency and sharing

“Privacy, what an abnormality”

Vint Cerf, one of the funders of internet

And transparency is ok.
But not at all costs.

Most of campaigners for total access to data have a financial interest in accessing data

e.g. medical journals and journalists: more articles

e.g. commercial companies, google…
Fundamental paradigm shift?

Figure 2: Heatmap visualization of disclosure patterns in the Carnegie Mellon Yearly Snapshot Dataset. As indicated by the legend, red colors indicate higher proportions of disclosure, and blues indicate lower proportions of disclosure. All variables are effects-coded at the unit level to capture if the participant shares the specific element.

Data from a panel of 5,076 Facebook users, first study to use data from Facebook's early days in 2005. Silent Listeners: The Evolution of Privacy and Disclosure on Facebook
Fred Stutzman. Journal of Privacy and Confidentiality (2012) 4, Number 2, 7-41
In the consent letter of the trial I’m in:

• I consented to the following:
  – Trial data (medical data, life style, socio-demography, sexual life, ethnic origin) will be transferred to:
    • The trial sponsor and persons acting on its behalf, in France or abroad
    • The French or foreign health authorities (Medicines Agency).
• No mention of anyone else who could access “my” data even in an anonymised way
• The term anonymised is not used, but “conditions that ensure data confidentiality”
• De-identification is not mentioned either
More importantly

• As part of the trial, your personal data will be processed to help analysing the trial results, in relation to the research objectives which were explained to you.

• No mention of a secondary use of the data
• No mention of a different use of the data than for the research objectives
• If properly explained: I would probably have agreed to sharing data with 3rd parties
Respect the consent form

• Current and future CTs: revise the way data sharing is mentioned in consent forms

• CT from the past
  – Not always possible to ask volunteers to re-consent
  – Some argue that doesn’t matter

• When the effort was made
  – 99% of volunteers re-consented

• Think of the 1% who refused, if their data would have been transferred anyway
  – a small but vociferous minority can damage trust in clinical research
CT participants don’t trust investigators 100%

• Regular press coverage of trials which went wrong
  – Hidden information
  – GCP issues when inspecting sites
  – CT which were clearly unethical (Tuskegee syphilis trial)

• Agreement between patients’ organisations and the National Aids Research agency, France
  – Information materials contain not only the volume of blood taken at each visit
  – But also the size and number of tubes
  – This is to prevent researchers from collecting blood for other purposes than the research in question
The inner circle: access to anonymised data

Patient
ID_data

Patient
data
ID_data

Anonymised
data

Study team
Investigator(s)
Nurse(s)

Statisticians

Sponsor

DSMB
The outer circle: access to de-identified data

- Anonymised data
- Sponsor
- Anonymised data
- Regulators
- De-identified data
  - Research organisations
  - Medical journals
  - HTA agencies
  - Competitors

Clinician
Sharing anonymised data with no control: the risks we see

• Identification of the patients possible (24-35%)
  – The *Unique in the Crowd* authors point to one case where a medical database was analysed against a voter list to discover the governor’s of Massachusetts health records
  – Latanya Sweeney, Carnegie Mellon Univ., showed that using 1990 census data and proving that 87 percent of Americans could be identified using just their ZIP code, sex and date of birth.

• Perhaps the best bet is try to ensure that valuable data, anonymous or not, doesn’t get into the wrong hands.

Data privacy

- **Age:** 50-year-old
- **Gender:** male
- **Suspected ADR:** respiratory arrest
- **Suspect drugs:** Dipriva, pethidine, alprazolam & sertraline
- **PMH:** low BMI, vitiligo, lupus syndrome
- **Outcome:** fatal
- **Occupation:** Rock star
- **Narrative:** patient was being treated by his personal physician at his mansion. Administration of propofol led to respiratory arrest and paramedics were called to assist. The resulting court case saw the attending doctor found guilty of involuntary manslaughter
Data privacy (2)

- **Age**: 41-year-old
- **Gender**: male
- **Suspected ADR**: Possible drug interaction with alcohol
- **Suspect drugs**: Fluoxetine, tiapride, and albendazole
- **PMH**: alcoholism?
- **Outcome**: fatal
- **Occupation**: Deputy Head of Security, Ritz Hotel, Paris
- **Narrative**: Patient was driving a Mercedes at high speed in the Pont d’Alma tunnel in Paris. He lost control of the car and crashed, killing himself and two passengers
Within the electronic record is private patient information including: physician diagnoses, procedures and payment information.

Raymond E. Boylston, the man identified in the news article, is linked to “anonymous record” #502855338.

MEDICAL RECORD

RECORD: 502855338
admitend: 10/25/11
STAYTYPE: 1
HOSPITAL: 162
COUNTYRES: 13
AGE_MONTH: 725
ZIPCODE 98851

Charges: $71708.47
DescDIAG1
80843: closed fracture of other specified part of pelvis; pelv fx-clos/pelv disrupt
DescDIAG2
5185: pulmonary insufficiency following trauma & surgery; post traum pulm insuffic
DescDIAG3
86500: injury to spleen without mention of open wound int
Re-identification with de-identified data: very low risk

• A Systematic Review of Re-Identification Attacks on Health Data
  – 6/14 re-identification attacks involved health data
  – Of which 2 studies where the original data was de-identified using current standards
  – Only one of these attacks was on health data, and the percentage of records re-identified was 0.013% (2/15 000), which would be considered a very low success rate

• [http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0028071#s2](http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0028071#s2)
Genetic information even more sensitive

• Researchers armed with little more than an internet connection identified nearly 50 people who participated in a large genomic study based on some of the participants’ genomes and other publicly accessible information
  – Linking with commercial genealogy databases. Yaniv Erlich of the Whitehead Institute for Biomedical Research
The privacy of our medical records is being sold off

- Genetic information can end up in the wrong hands. It could result in anything:
  - from discrimination by unscrupulous employers
  - or insurance companies
  - or for DNA being synthesised and planted at a crime scene

- Business access to private data: already happening:
  - USA: states selling their health data
  - France: Saint Malo hospital granting access to confidential data to a service provider (www.syndicat-infirmier.com/Hopitaux-des-infractions-aux.html)
  - GP and hospital records made available to drug company and other researchers through the Clinical Practice Research Database (www.theguardian.com/commentisfree/2012/aug/28/code-practice-medical-data-vulnerable)
Database linkage and privacy protection

• “Anonymity is a myth if you’ve got richly detailed genetic information and access to a variety of databases”

• “Researchers need to ensure informed consent from participants even if that means telling them it may not be possible to protect their privacy”
  – Hank Greely, a law professor at Stanford University who specialises in the ethical and legal implications of emerging biotechnology
There is a greater risk than re-identification

- Secondary analysis of same data can generate different results
  - Patients who were first excluded now included
  - Different analytic method, standards applied or not
- If different results are obtained: need for a scientific dialogue to understand why, before going public
- If new results go public: risk of unnecessary controversy
- We don’t need to end up in the same situation than for GMO policy
- The consequence would be volunteers running away from clinical trials
What we propose:

- Adapt consent letter: the text is never too clear
- If continuity with initial research: no need to re-consent
- No risk of “consent fatigue”, rather “dynamic consent”
- Researchers who access data for secondary use
  - Default policy: de-identified data
  - If need for anonymised or identifiable data: consent requested, ECs opinion
  - Before publishing their own results, if different from initial ones: commitment to first discuss their results with initial team or regulators
- Institutions/researchers who would not respect this commitment: black list, and exception in consent letters
Access to de-identified or anonymised data: main issues

- Secondary analysis of data and consent
- Risks of identifying people
- Analytic utility of severely modified data
- Controversy & inadequate secondary analysis
- Publication of methods and results / review
The main challenge: to make clinical research more **attractive**

A new concern: will access to data generate **knowledge**? Or **noise**?

We like it or not, our data, the **new petrol**, belong to business

Maybe data protection already belongs to the past.
If there is a need to access back to the patient or his data

- Any third party researcher (out of the “inner circle”) can contact the study team and request access to the patients
  - Submit a research proposal
    - The purpose of the data access request is explained to the patients
    - In some cases a re-consent will be needed
- In all cases, the “inner circle” will always have access to the patients (via the access key)
- No need to release anonymised data to third parties
No mention of a “transfer to third parties”, but:

• If purpose properly explained: I would probably have agreed to sharing de-identified data to third parties
  – Even if third party not known at that time

• But as I gave my consent for my anonymised data to be only accessed by a restricted group of people, I would feel betrayed if they would be transferred to a third party without my consent
Who’s behind the re-identification attacks?

- 14 attacks analysed 2001-2010
- 11 by researchers
  - To demonstrate that a risk exists or to evaluate if one exists, but not to exploit that risk
- 1 by a journalist
- 1 by an expert witness
  - To inform a court judgment
- 1 by a broadcaster
  - To inform a court judgment
The informed consent, corner stone of trial participants’ protection

• The process of obtaining the consent from trial participants is what distinguishes volunteers from guinea pigs

• We give our data and tissues to research, in the interest of science and society

• However, the Helsinki declaration clearly states
  – “Concern for the interests of the subject must always prevail over the interests of science and society”
The case of international registries

In each country, data are collected and stored in the "inner circle". Difficulties arise when all data need to be merged in a single database: data transfer to another country extremely difficult.

The concept of "inner circle" or study team should be expanded to the "registry team".

Data transfer to a single EU database should be made easier.