Appendix:

Science leaks: a signal to improve data protection in scientific research

Published as:
van der Aa HPA, van Nispen RMA, Kliphuis T, Paardekooper M, Knol D, van Rens GHMB. Science leaks: a signal to improve data protection in scientific research
J Clinic Res Bioeth 2015; 6:221. DOI: 10.4172/2155-9627.1000221

Threat of privacy invasion

"Hello sir, I'm calling regarding a recent scientific study in which you are participating. I want to thank you for taking part in this study by offering you a present. Only a small amount of postage costs need to be paid to receive this present, for which I will need your bank account number."

Recently, an older visually impaired participant of a clinical trial conducted by the VU University Medical Centre in Amsterdam (which will be referred to as the university hospital in this article) received this harassing telephone message from someone unknown to the research team. Fortunately, the participant did not respond to the request and immediately contacted the executive researcher.

This incident is a textbook example of 'phishing', i.e. the act of attempting to acquire information from an individual by masquerading as being trustworthy via electronic communication. Phishing is an example of social engineering, i.e. psychological manipulation with the aim to acquire confidential information.[1,2] The threat of privacy invasions is progressing and technological advances in the digital world play a significant role contributing to this trend. In recent years reports about spying, wiretapping and invasion of privacy frequently dominated the media and provoked a worldwide debate on the importance of privacy.

In scientific research, the right to privacy versus the responsibility of society to conduct valuable medical research is an ongoing issue of debate.[3-7] A major topic in this debate is that of 'nourishing' the patient's trust. Medical research would not be possible without the patient's trust: when patients are not ensured and convinced that their private information is handled confidentially, they may not be inclined to participate.[3,6] Instructions on how to prevent privacy violations within clinical research have been published, e.g. not discussing patients in the public area and following strict security protocols when working with electronic data.[8] However, even when professionals follow these rules and regulations carefully, the threat of privacy invasions, e.g. the phishing incident described in this article, cannot always be completely eliminated. This threat is increasing because of technological progression, growing complexity of large clinical trials, growing amounts of electronic datasets and the increasing value of personal medical data on the black market.[9] This raises the question: how can data security be improved to reduce the increasing threat of privacy invasions? In this commentary article we describe measures to deal with the phishing incident and to improve data protection, supported by National and European regulations.

Approaching privacy invasion

To approach the described phishing incident at the university hospital we undertook three main steps, i.e. 1) reporting the incident, 2) dealing with the incident, and 3) follow-up. These steps will be explained in the following sections.

1. Reporting the incident

1.1 Discovering and reporting the privacy violation

The harassed participant reported the incident by email; this was immediately recognised by the executive researcher (HvdA) and the principal investigator (RvN) as an intrusion of privacy. That same day, the principal investigator contacted the data protection officer of the university hospital (TK) to report the incident and discuss how to deal with the situation.

1.2 Initial context analysis

The data protection officer responded immediately and initiated a context analysis based on: 1) contacting the participant to obtain more details, 2) investigating possible data leaks within the research team/setting, and 3) investigating whether other participants of the clinical trial were approached, in order to assess the impact and determine appropriate steps. In addition, the board of directors of the university hospital was immediately informed and monitored the entire procedure.

2. Dealing with the incident

2.1 Contacting the harassed participant

Being a patient of a low vision rehabilitation centre, the harassed participant had provided written informed consent to participate in the clinical trial performed at the university hospital (*n*=908). The participant was contacted by the executive researcher to offer support and acquire more details on the incident. Fortunately, the participant did not experience any negative consequences. The telephone number of the intruder could not be obtained.

2.2 Investigating possible data leaks

Processing information in large clinical trials is complex. A large amount of information needs to be processed, entailing complex data storage and the involvement of multiple professionals. Various sources of a possible data leak could have occurred within the university hospital via: 1) the informed consent forms: these were kept in a file on an open shelf in the executive researcher's office (which was always locked when not occupied); 2) information kept on a personal network account protected by a variable password; 3) six research assistants working in shifts to interview the study participants. Other sources of a data leak could have arisen outside the university hospital: 4) the low vision rehabilitation centre where all eligible participants were registered; 5) the general practitioner who was notified; 6) three sealed envelopes with the logo of the university hospital that were sent to the harassed participant's home address; 7) various 'significant others' who also knew about the clinical trial, i.e. the participant's wife, various friends, and his ophthalmologist; and 8) any random person who may have overheard 'revealing' conversation(s). No data were found missing or destroyed. However, after contacting the various organisations and people involved, the possibility that a leak had occurred in any of these places could not be totally excluded.

2.3 Investigating if other participants were approached

It was not feasible to contact all participants of the clinical trial to determine if they were also approached by the intruder. Therefore, a randomly selected sample of participants was contacted by telephone as part of a regular call made for study purposes. To determine an appropriate sample size of this random selection, the level of certainty of finding other participants that were harassed by the intruder was determined.

The hypergeometric distribution was used that describes the probability of the number of harassed participants in n (sample size) without replacement from a finite population of size n containing exactly M defects.10 Numbers of additionally harassed participants were assumed from the lowest probability (M=1) to the higher probability (M=40). The probability of finding one other incident in a sample of 80 participants (8% of the study population), under the assumption that one other participant was approached (M=1) by the intruder, was 8.8%. The probability of finding one or more incidents in this sample under the assumption that the intruder had called 20 additional participants was 84.5% and for 30 participants this was 94.0% (see TABLE 1). A sample size of 80 participants was chosen as this was a feasible number of participants to call and ruled out the possibility that 20 or more additional participants were harassed with great certainty (more than 80%). Based on this sample no other participants were found that had been harassed by the intruder (95% confidence set for M: {0,0…,31}.11 Therefore, the possibility that the intruder had gained access to the data of all participants was low. Furthermore, no similar incidents were reported by any other study participant of the clinical trial. Had that been the case, more spontaneous notifications would have been expected.

TABLE 1. Hypergeometric probabilities of finding more harassed participants for various numbers of M, for a finite population size (***n***=908)

No. of harassed persons in the study population (M)	Sample size (n)	Probability (0)	Probability (≥1)
1	80	0.912	0.088
20	80	0.155	0.845
30	80	0.060	0.940
40	80	0.023	0.977

3. Follow-up

After the investigation of the incident was completed, the board of directors of the university hospital, the board of directors of the low vision rehabilitation centres, as well as the harassed participant himself, were informed about the incident and the steps that were undertaken. Since this was a relatively minor incident (with limited effect on the participant's wellbeing) it was not necessary for the data protection officer to inform the data protection authority

Prevent incidents in the future

The incident described here illustrates that phishing and other types of data leaks and/or security breaches can be a threat in scientific research. National and European laws provide a framework to prevent and deal with these situations and stimulate uniformity and precision (see box 1 for a brief overview). The Medical Research Involving Human Subjects Act (as part of the Declaration of Helsinki) states that when interventional research in human subjects is carried out, both an ethics committee and the government should authorize the execution.[12] In addition, the Dutch Data Protection Act states that researchers should ensure that the privacy of participants is

protected whenever possible. It is expected that more incidents will be revealed when this act is updated (planned during 2015): in case of violation of privacy and/or security breaches, the data protection authority must be notified within 24 hours. In England and some other European countries notification of data breaches is already obligatory. The English data commissioner office published data breach trends showing about 50% of reported incidents in the health domain in the last quarter of 2014.[13]

These national laws follow the European Data Protection Directive 95/96, which states that anyone holding personal data (other than for domestic use) is legally obliged to comply with this act.5 However, the European Data Protection Directive does not sufficiently consider important aspects such as globalisation and technological developments. Therefore, the General Data Protection Regulation is planned to supersede this latter directive (expected in 2015-2016). The General Data Protection Regulation aims to unify data protection within the European Union, including severe penalties for security breaches and non-compliance with the regulation. The regulation will focus on aspects such as social networks and cloud computing.[14]

BOX 1: Rules and regulations concerning research on human subjects

National (Dutch)	
Medical Research Involving Human Subjects Act (Established in 1998, changed in 2006)	- Ethics committee and government should authorize execution. - Study protocol, extensive information letter and informed consent form should be written and documented according to certain standards. - Independent physician should be available to provide additional information for participants.
Data Protection Act (Established in 2001)	- The data protection authority must be notified of all processing of personal data. - Organisations can appoint their own internal supervisor, the data protection officer. - Subjects have the right to know what happens to their personal information and to access their personal data. - Valid consent must be explicit for collected data.
European	
European Data Protection Directive 95/96 (Established in 1995)	Seven principles are defined: 1. Subjects should be given notice when their data are collected; 2. Data should only be used for the stated purpose; 3. Subjects should always give consent; 4. Data should be kept secure from any potential abuses; 5. Subjects should be informed about who is collecting their data; 6. Subjects should be able to access and make corrections to their data. 7. Subjects should be able to hold data collectors accountable for adhering to all of these principles.
General Data Protection Regulation (Expected in 2015-2016)	- Data protection of all European Union residents (even concerning foreign companies). - Same rules apply to all European Union member states. - Valid consent must be explicit for collected data. - Rights of subjects are enhanced. - Data Protection Impact Assessments are obligatory. - Organisations are forced to prove compliance to the regulation - The data protection authority needs to be notified within 24 hours after having become aware of the data breach (when feasible). - Subjects have to be notified if adverse impact is plausible. - Sanctions can be imposed up to a fine of 100 million Euros or up to 5% of the annual worldwide turnover.

Each organisation has to find a way to implement uniform protocols to secure and improve the patients' privacy; taking legal demands, demands of research funders on data sharing and increased (international) collaboration into account. This is not the preserve of a few individuals but needs to involve the entire academic and research field. However, the interpretation of these regulations is not always clear and there is a dearth of clear policy guidance. The steps that were taken in this article may help researchers to deal with events of privacy intrusions. Crucial elements in this process proved to be: recognising the incident as a privacy violation, acting immediately after the incident had occurred, systematically retracing possible data leaks, and involving the victim during the process.

It is of pivotal importance to increase awareness of the threat privacy invasions and to provide uniform protocols for all employees to be able to detect and act immediately to deal with such violations. Data protection officers, privacy officers or other employees who are responsible for providing guidance to researchers in such situations should be easily reachable within the organisation to enable researchers to act quickly. In addition, patients can be important 'partners' in fighting privacy violations.8 Therefore, it is recommended that participants of clinical trials be informed of possible threats of privacy invasions and instructed to report such incidents to the researchers should they occur. Information must be provided carefully, in such a way that distrust is avoided and willingness to participate in the studies enhanced. This can be achieved by providing an information letter which may state that i) researchers will never ask for information other than that directly necessary for the study, and ii) that participants should always verify the telephone number and/or e-mail addresses used to contact them in relation to the study.

Proper precautions must be taken to restrict the chance of privacy violations to a minimum, and to facilitate adequate response should such incidents occur, to maintain the trust of participants within scientific research. Only when individuals are convinced that their personal data are being used under strictly controlled conditions, are they likely to agree to offer up some individual privacy for the greater societal good that can emerge from scientific research.

Key Messages

Awareness of the growing threat of privacy invasions when working with patients within scientific research settings should be increased. The steps that were taken after the phishing incident described in this article might serve as a guideline to deal with privacy violations, in which crucial elements proved to be: recognising the incident as an intrusion of privacy, acting immediately after the incident had occurred, systematically retracing possible data leaks, and involving the victim in the process. Uniform privacy protocols should be provided within the academic research field to deal with these situations. New legislation (national: Data Protection Act, European: General Data Protection Regulation) offer a framework to prevent and deal with these situations.

References

1. McCombs B. Phoney phishing and pharming. Can J Rural Med 2005; 10:186-7.
2. Mayhorn CB, Nyeste PG. Training users to counteract phishing. Work 2012; 41:3549-52.
3. Taylor MJ. Health research, data protection, and the public interest in notification. Medical Law Review 2011; 19:267-303.
4. Walley T. Using personal health information in medical research. BMJ 2006; 332:130-1.
5. Iversen A, Liddell K, Fear N, Hotopf M, Wessely S. Consent, confidentiality, and the Data Protection Act. BMJ 2006; 332:165-9.
6. Coleman MP, Evans BG, Barrett G. Confidentiality and the public interest in medical research – will we ever get it right?. Clin Med 2003; 3:219-28.
7. Strobl J, Cave E, Walley T. Data protection legislation: interpretation and barriers to research. BMJ 2000; 321:890-2.
8. Taitsman JK, Grimm CM, Agrawal S. Protecting patient privacy and data security. New England Journal of Medicine 2013; 368:977-9.
9. EMC Corporation, RSA security. Cybercrime and the healthcare industry. http://www.emc.com/collateral/white-papers/h12105-cybercrime-healthcare-industry-rsa-wp.pdf (accessed 20 January 2015).
10. Rice JA. Mathematical Statistics and Data Analysis (Third edition). Third edition Duxbury, United States: Duxbury Press, 2007:42.
11. Blaker H. Confidence curves and improved exact confidence intervals for discrete distributions. Canadian Journal of Statistics 2000; 28:783-98.
12. World Medical Association. World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. Bull World Health Organ 2001; 79:373–374.
13. Information Commissioner's Office. Data breach trends. https://ico.org.uk/action-weve-taken/data-breach-trends (accessed 20 January 2015).
14. De Hert P, Papakonstantinou V. The proposed data protection Regulation replacing Directive 95/46/EC: A sound system for the protection of individuals. Computer law and security review 2012;,30:130-42.