Published version available: Rumbold, John Mark Michael and Pierscionek, Barbara (2017) The effect of the General Data Protection Regulation on medical research. Journal of Medical Internet Research, 19(2), e47. at https://www.jmir.org/2017/2/e47/

Implications of the General Data Protection Regulation for medical research

John Rumbold*, Barbara Pierscionek

a) Postdoctoral Fellow b) Associate Dean Research & Enterprise Faculty of Science, Engineering and Computing, Kingston University London, Penrhyn Road, Kingston upon Thames KT1 2EE UK

Corresponding author: John Rumbold

*Email: J.Rumbold@Kingston.ac.uk

Abstract

Background: The enactment of the General Data Protection Regulation (GDPR) will impact on European data science. Particular concerns relating to consent requirements that would severely restrict medical data research have been raised.

Objectives: To explain the changes in data protection laws that apply to medical research and to discuss their potential impact.

Methods: Analysis of ethico-legal requirements imposed by the GDPR.

Results: The GDPR makes the classification of pseudonymised data as personal data clearer, although it has not been entirely resolved. Biomedical research on personal data where consent has not been obtained must be of substantial public interest.

Conclusions: The GDPR introduces protections for data subjects that aim for consistency across the EU. The proposed changes will make little impact on biomedical data research.

Keywords: Pseudonymity, anonymity and untraceability; privacy-preserving protocols; informatics; data reporting; data protection; research ethics

There have been significant developments in European Union (EU) data protection law recently that will have an impact on healthcare professionals particularly those engaged in research and audit. The General Data Protection Regulation (GDPR) has replaced the current legislation and comes into full effect in 2018. [1] The implications for the handling of healthcare data of the GDPR will be discussed in this paper. Despite the recent referendum vote in the UK to leave the EU, the GDPR will continue to be relevant to the UK, whether this is due to cooperation in European projects or because the UK continues to be a member of the European Economic Area (EEA).

The Data Protection Directive

Currently the relevant law in the UK is the Data Protection Act 1998, which is the UK's transposition of the Data Protection Directive (DPD). European directives are not directly enforceable, requiring member states to pass legislation to comply with their requirements. There are derogations (legal exemptions) for research, which in the case of the UK have been criticized for being too broad. The LRDP Kantor report

for the European Commission criticises the UK for disregard of the limitations, stating that the Data Protection Act blatantly violates the Directive by adding “medical research” to the list of medical purposes. [2] The DPD requires a “substantial public interest” for member states to add to the derogations for processing of sensitive personal data (Article 8.4).

Differences between EU member states can result in research ethics committees in UK denying permission for NHS data to be transferred to other EU countries (the opposite might also be the case in some circumstances).[3] These differences have also contributed to the passage of the GDPR as part of the Digital Single Market strategy.[4]

The law as it will be from 2018: the GDPR

The text of the GDPR has recently been agreed after a prolonged trilogue between the European Commission, Parliament and Council of Ministers. [5] This legislation will replace the national transpositions of the DPD. Regulations are directly enforceable across the EU. The GDPR comes into full effect on May 25th 2018, although member states are permitted minor differences in interpretation (the European Court of Justice is the ultimate arbiter). This legislation has the potential to affect projects using research data banks and/or Big Data.[6, 7] There had been concerns that a clause inserted by the European Parliament requiring specific consent would prevent significant long-term epidemiological research taking place in the future,[8]but this was rejected and the agreed text permits broad consent to “certain areas of research when in keeping with recognised ethical standards” (Recital 33).[9] Broad consent is not blanket or open consent [10] although some commentators argue that blanket or open consent is acceptable for biobank and databank research as the risks are minimal and do not vary for different projects.[11] Another possibility is consent to a form of governance.[12] Open consent without any ongoing regulation or communication about proposed projects would be potentially problematic. Dynamic consent offers advantages for an engaged community of participants but might not be considered beneficial by some individuals.[13]

The derogations for research without consent have been expanded to specifically include medical research where “in the public interest” (Recital 51). How public interest will be defined has not been elaborated, but European jurisprudence demands member states satisfy a high threshold where human rights are involved e.g. a “pressing social need”.[14] This standard would not be required for the conduct of medical research using databanks, but it might exclude all commercial research for “me too” drug development (drugs that offer no advantages over drugs already on the market), arrangements that have no evidence of benefit sharing, or simply require that projects address issues of public importance, regardless of the profits made.[15] This requirement reflects public attitudes in the UK to the use of healthcare data, where there is resistance to use of public data for commercial ventures unless the research could not happen without commercial involvement.[16, 17]

Anonymisation

Data protection law only applies to personal data – that is, data that does directly or can indirectly identify an individual.[18-20] The simple deletion of name and address is usually insufficient to constitute anonymisation (it has been demonstrated that the combination of three pieces of data could identify 87% of US residents – five-digit zip code, birth date, and sex).[21] The UK Information Commissioner's Office currently treats pseudonymised data as anonymous where it is used by a third party who does not possess the requisite key code. Truly anonymised data cannot be linked back to an individual (which means that verification of data is not possible by any means). Pseudonymised data typically has identifiers removed and replaced with a unique key code (there is also two-way cryptography; one-way cryptography is considered anonymised). This key code can be used to trace the data back to an individual, enabling any safety concerns to be acted upon and for data to be verified. This is the approach that the UK Care.data project on the use of NHS electronic health records for data research has been taking [22]. The GDPR will require changes in practice, as it confirms in Recital 26 that pseudonymised data must be treated as personal data (in line with the previous Article 29 Working Party opinion).[18] That position results from the increased vulnerability of data subjects who could potentially be identified compared to the protection afforded them with true anonymisation - if the key code is hacked, then all the data can be linked to an individual once more.

Consent

Consent presumed by failure to "opt-out" or change pre-ticked boxes will no longer be permitted (unless covered by the derogations) – consent will need to be by a "clear, affirmative action" (Article 4.11). These changes would have arguably made the abandoned Care.data project[23] illegal, despite the passage of enabling legislation that exempted general practitioners from the common law duty of confidentiality when fulfilling their contractual duties to pass on healthcare data. Care.data relied on an opt-out for legitimacy[22]. The exercise of this opt-out was not straightforward. The numbers opting out far exceeded the estimates and the capacity for the Health and Social Care Information Centre (now NHS Digital) to process in a timely manner. The problems included omission of those who opted out from calls for NHS screening programmes, even though this was not the intention of those exercising this right. NHS Digital currently relies on pseudonymisation, which the GDPR states is categorized as a matter of law as personal data. It is not entirely clear whether or not third parties without access to the key code could treat pseudonymised data as anonymised (as is currently the case in the UK). Key codes are a potential vulnerability due to accidental or malicious disclosure, which is one of the justifications for pseudonymised data being classified as personal data. There are no clear indications that there are no future plans to use NHS patient data for research.

Dame Fiona Caldicott reviewed arrangements because of the widespread concerns related to consent;[22] and her report led to the cancellation of the Care.data project.[23] The particular issues that were identified include the lack of information about Care.data that made exercising an opt-out an opaque process, the inadequate mechanisms for opting, and the failure of protection for rights and access to the NHS for those who opt out.

The risk of re-identification in the future is impossible to quantify precisely, because it cannot be predicted what information will become public.[24] However, as with biobanks, the risks to individuals are lesser compared with studies of medical interventions.[25] Therefore authorization by research ethics committees is acceptable practice, with the requirement that opt-outs be respected unless there are exceptional circumstances.

Although the GDPR comes into force in mid-2018, researchers need to prepare now for the changes it will bring to long-term epidemiological studies. In particular, the categorisation of pseudonymised data as personal will require action in some jurisdictions such as the UK and Greece.[26] The necessary accommodations will require an investment of resources, but this will hopefully ensure that subjects continue to have trust in the integrity of their healthcare data and the medical research community.[27] The GDPR may still apply should the UK cease to become a member state of the EU either because the UK is a member of the EEA or because the UK retains these instruments as law at least for the short term.[28]

Although audit and research are treated differently in law, the boundaries between the two activities are blurred.[29] Audit is directly relevant to the monitoring and improvement of quality of healthcare; therefore, it is included as a primary use of data (Recitals 52-54 and Article 9.2 (h) and (i) of the GDPR make this clear). Audit and healthcare management are a primary use of healthcare data, and research is a secondary use – that is, it is a use different from the originally declared purpose (although it is designated a compatible purpose within the GDPR, but only for non-sensitive data). If an audit compares healthcare systems to discover which is most effective, this can also be categorized as research as the practices are not compared to a gold standard, and there is a hypothesis being generated or even tested by finding associations. The recent furore surrounding the Royal Free Trust project in conjunction with Google DeepMind illustrates the debate over the distinction of audit from research.[30-32]

Data Sharing

Dame Fiona Caldicott affirmed in her 2013 report on information governance that

> “The duty to share can be as important as the duty to protect patient confidentiality.”[33]

Data sharing within the EU should not be obstructed because of differences in data protection law, under the principles of the Digital Single Market and Article 1(2) of the Data Protection Directive. Data portability and data sharing is an issue with healthcare data,[34] which the epSOS project attempted to address.[35] The GDPR addresses data portability under Article 20, stating that the data subject has the right to receive their data in an appropriate format without hindrance and for data to be transferred between data controllers where technically feasible. The Bundestag is currently considering an eHealth bill with the same aim of improving portability of data.[36] This will facilitate the ability of patients to move between healthcare providers without unnecessary duplication of tests.

Conclusions

The Digital Single Market aims for improved data sharing across the EU which will facilitate cross-border healthcare and research. Harmonisation will be improved under the GDPR with a concomitant raising of standards for some countries, although there is still room for national differences according to the reasonable expectations of different publics. This advance makes cross-border projects more easily ethically justifiable and more feasible.[37] The requirements for anonymisation have not been changed, except to clarify that pseudonymised data must still be considered as personal data. The GDPR will facilitate medical research, *except* where it is research not considered in the public interest. In that case, more demanding requirements for anonymisation will entail either true anonymisation or consent. It is likely there will be more projects that require either consent or authorisation, since many projects currently use pseudonymisation. There is still an unresolved issue over third parties with access to pseudonymised data.

Footnotes

Contributors: Both authors contributed to the analysis of legal issues and the writing of the manuscript.

Funding: This work has been partly funded by AEGLE project, Horizon 2020 ICT/2014/1 grant

Conflicts of interests: We have nothing to declare

References

1. European Union. General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679). Official Journal of the European Journal. 2016 May 4th(L119).

2. LRDP Kantor in association with the Centre for Public Reform. New Challenges to Data Protection, prepared for the European Union's Directorate-General Freedom Security, Justice (2010): http://ec.europa.eu/justice/policies/privacy/docs/studies/new_privacy_challenges/final_report_en.pdf

3. Veerus P, Lexchin J, Hemminki E. Legislative regulation and ethical governance of medical research in different European Union countries. Journal of Medical Ethics. 2013 May 10:medethics-2012.

4. Reform of Data Protection Rules (2016) [Internet] Brussels: European Commission [updated Jun 2016]. Available from: http://ec.europa.eu/justice/data-protection/reform/index_en.htm

5. Interinstitutional File: 2012/0011 (COD) [Internet].; 2015 []. Available from: Statement by Vice-President Andrus Ansip at the press conference on the adoption of the Digital Single Market Strategy.

6. The 5 V's of Big Data [Internet].: Data Science Central; 2015 [updated Apr 9th]. Available from: http://www.datasciencecentral.com/profiles/blogs/the-5-v-s-of-big-data-by-bernard-marr.

7. Analysis: Research and the General Data Protection Regulation - 2012/0011(COD) [Internet]. London: Wellcome Trust [updated July2016; cited Dec 2nd 2016]. Available from: https://wellcome.ac.uk/sites/default/files/new-data-protection-regulation-key-clauses-wellcome-jul16.pdf.

8. Stevens L. The Proposed Data Protection Regulation and Its Potential Impact on Social Sciences Research in the UK. European Data Protection Law Review. 2015;1(2):97-112.

9. Simon CM, L'Heureux J, Murray JC, Winokur P, Weiner G, Newbury E, et al. Active choice but not too active: Public perspectives on biobank consent models. Genetics in Medicine. 2011;13(9):821-31.

10. Hofmann B. Broadening consent - and diluting ethics? J Med Ethics. 2009;35:129-29.

11. Sheehan M. Can Broad Consent be Informed Consent? Public Health Ethics. 2011;4(3):226-35.

12. Laurie G. Governing the Spaces In-Between: Law and Legitimacy in New Health Technologies. In: Flear ML, Farrell A, Hervey TK, Murphy T, editors. *European Law and New Health Technologies*. Oxford: Oxford University Press; 2013. p. 193.

13. Steinbekk KS, Myskja BK, Solberg B. Broad consent versus dynamic consent in biobank research: Is passive participation an ethical problem? European Journal of Human Genetics Open. 2013;21:897-902.

14. *Handyside v United Kingdom*. EHRR. [1976];1:737.

15. Haddow G, Laurie G, Cunningham-Burley S, Hunter KG. Tackling community concerns about commercialisation and genetic research: A modest interdisciplinary proposal. Social Science & Medicine. 2007;64:272-82.

16. Aitken M. SHIP Public Engagement: Summary of Focus Group Findings. Scottish Health Informatics Programme; 2011.

17. Ipsos MORI. The One-Way Mirror: Public attitudes to commercial access to health data. London: Wellcome Trust; 2016.

18. Article 29 Data Protection Working Party. Opinion 4/2007 on the concept of personal data. Opinion. Brussels: Directorate C, European Commission; 2007. Report No.: 01248/07/EN WP136.

19. *R v Department of Health, ex parte Source Informatics*. QB. [2001]:424.

20. *Common Services Agency v Scottish Information Commissioner*. UKHL. [2008]:47.

21. Sweeney L. Simple Demographics Often Identify People Uniquely. Pittsburgh: Carnegie Mellon University; 2000. Report No.: Data Privacy Working Paper 3.

22. Caldicott: care.data hangs on engagement [Internet].; 2015 []. Available from: http://www.digitalhealth.net/analytics/46830/caldicott:-care.data-hangs-on-engagement.

23. NHS England to close care.data programme following Caldicott Review [Internet].: National Health Executive; 2016 [updated July 7th; cited July 29th 2016]. Available from:

http://www.nationalhealthexecutive.com/Health-Care-News/nhs-england-to-close-caredata-programme-following-caldicott-review.

24. Information Commissioners Office. Anonymisation: managing data protection risk code of practice. London: ICO; 2012.

25. Laurie G, Jones KH, Stevens L, Dobbs C. A Review of Evidence Relating to Harm Resulting from Uses of Health and Biomedical Data. Scoping Study. 2014.

26. European Forum for Good Clinical Practice. Data protection and research in the European Union (2015). Available at: http://www.efgcp.eu/downloads/DP%20and%20Research%20in%20EU_HD_Final_06%2010%2015.pdf

27. Carter P, Laurie GT, Dixon-Woods M. The social licence for research: why *care.data* ran into trouble. Journal of Medical Ethics. 2015:pp.medethics-2014.

28. Mason R. Theresa May's 'great repeal bill': what's going to happen and when? The Guardian. 2016 Oct 2nd;Sect. Politics.

29. Wade DT. Ethics, audit, and research: all shades of grey. British Medical Journal. 2005;330:468-73.

30. Hodson H. Google knows your ills. New Scientist. 2016;230(3072):22-23.

31. Shah NR, Seger AC, Seger DL, Fiskio JM, Kuperman GJ, Blumenfeld B, et al. Improving acceptance of computerized prescribing alerts in ambulatory care. Journal of American Medical Informatics Association. 2006;13(1):5-11.

32. ICO probes Google DeepMind patient data-sharing deal with NHS Hospital Trust [Internet].: ComputerWeekly.com; 2016 [updated May 12th; cited Jul 28th 2016]. Available from: http://www.computerweekly.com/news/450296175/ICO-probes-Google-DeepMind-patient-data-sharing-deal-with-NHS-Hospital-Trust.

33. Caldicott F, Independent Information Governance Oversight Panel. Information: To share or not to share? The Information Governance Review. London: Department of Health; 2013.

34. Kish LJ, Topol EJ. Unpatient - why patients should own their medical data. Nature Biotechnology. 2015;33(9):921-24.

35. Cross-border health project epSOS: What has it achieved? [Internet].: European Commission; 2014 [updated 07/07/2014;]. Available from: https://ec.europa.eu/digital-single-market/en/news/cross-border-health-project-epsos-what-has-it-achieved.

36. Act on secure digital communication and applications in the health care system (E-Health Act) [Internet].: Federal Ministry of Health; 2015 [updated Sep 29th;]. Available from: http://www.bmg.bund.de/en/health/e-health-act.html.

37. Dove ES, Townend D, Meslin EM, Bobrow M, Littler K, Nicol D, et al. RESEARCH ETHICS. Ethics review for international data-intensive research. Science. 2016 Mar 25;351(6280):1399-400.