厚生労働行政推進調査事業費補助金 (厚生労働科学特別研究事業)
「ゲノムデータの持つ個人識別性に関する研究」

分担研究報告書

欧米におけるゲノムデータの利用にかかる法制度 ―欧州データ保護指令/規則および米国 HIPAA プライヴァシー規則の匿名化ルールを中心に―

分担研究者　佐藤 智晶　　青山学院大学法学部　准教授

欧米において、いわゆる「ゲノムデータ」は、法制度上、日本でいうところの「個人識別符号」に該当するものと考えられており、その結果として一定の法的保護の対象とされているが、厳密にどのようなゲノムデータならば個人識別性を帯びるのかについては議論が深まっていない[1]。むしろ、ゲノムデータには個人識別性があるという前提のもとで、十分な匿名化を施せば本人同意なしに収集や利用ができる、というのが欧米の見解である[2]。

欧米を比べた場合、やはり欧州の方が匿名化（anonymisation）の要件は厳しく、ゲノムデータを同意なしで利用することは難しい。たとえば、欧州委員会 29 条作業部会の見解（Opinion 05/2014 on Anonymisation Techniques）によれば、ゲノムデータが個人識別性を帯びていることを前提に、提供元が個人特定できない処理を実施して作成したデータセットを第三者に提供し、第三者がデータセットを適切に使う場合（再特定しないで使う場合）、本人の同意なしにゲノムデータを利用する可能性を必ずしも除外していない[3]。しかしながら、十分な匿名化が実施されているか否かについては、提供元に残されている元データや他のデータセットだけでなく、提供先が参照可能な他のデータが考慮されるだけでなく[4]、さらに、個人特定にかかる費用と時間、匿名化処理を施した際

---

[1] 欧州に関する最新の論文としては、たとえば、次のものを参考にされたい。See Jane Kaye,et al., Med Law Rev (2013) doi: 10.1093/medlaw/fwt027 First published online: October 17, 2013, available at http://m.medlaw.oxfordjournals.org/content/early/2013/10/17/medlaw.fwt027.full

[2] Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, 0829/14/EN, WP216, 10 April, 2014, n. 27, available at http://www.cnpd.public.lu/fr/publications/groupe-art29/wp216_en.pdf

[3] Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, 0829/14/EN, WP216, 10 April, 2014, at 10, available at http://www.cnpd.public.lu/fr/publications/groupe-art29/wp216_en.pdf

[4] Id. at 9 (…"Thus, it is critical to understand that when a data controller does not delete the original (identifiable) data at event-level, and the data controller hands over part of this dataset (for example after removal or masking of identifiable data), the resulting dataset is still personal data. Only if the

の技術水準だけでなく今後の技術的な発展のようなすべての客観的な事情が考慮される[5]。今後の技術的な発展まで考慮して個人識別性を帯びるかどうかが判断されるとすれば、提供元としては相当に匿名化を施さなくてはならない。

米国でも、欧州と同じようにゲノムデータは原則として法的な保護対象であるが、十分な匿名化を施すことによって本人の同意なしに利用することができる[6]。HIPAA プライヴァシー規則が制定された当初から、匿名化の方法については賛否両論があったとされる[7]。なぜならば、HIPAA 法の規制対象となるのは、個人識別可能な医 療情報（individually identifiable health information）の取り扱いであったため、 規制対象外となる条件、すなわち、医療情報の適法な匿名化について関心が集まることになった。

先に説明したとおり、同規則における匿名化の方法が実際に変更されたことはないが、 最初の規則（案）からは一度だけ変更されて規則として公表されている。2000 年の 12 月 28 日に公表された規則と、1999 年 11 月 3 日に公表された同規則（案）では、明ら かに匿名化の方法が異なっている。規則（案）では、所定の情報を除去すれば匿名化された情報と推定する、という規定にな

---

data controller would aggregate the data to a level where the individual events are no longer identifiable, the resulting dataset can be qualified as anonymous. For example: if an organisation collects data on individual travel movements, the individual travel patterns at event level would still qualify as personal data for any party, as long as the data controller (or any other party) still has access to the original raw data, even if direct identifiers have been removed from the set provided to third parties. But if the data controller would delete the raw data, and only provide aggregate statistics to third parties on a high level, such as 'on Mondays on trajectory X there are 160% more passengers than on Tuesdays', that would qualify as anonymous data."

[5] Id. at n. 6. See also The Final version of the EU General Data Protection Regulation, Recital 23, Dec. 15, 2015, available at https://iapp.org/media/pdf/resource_center/2015_12_15-GDPR_final_outcome_trilogue_consolidated_text.pdf (..." To determine whether a person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by any other person to identify the individual directly or indirectly. To ascertain whether means are reasonable likely to be used to identify the individual, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration both available technology at the time of the processing and technological development. The principles of data protection should therefore not apply to anonymous information, that is information which does not relate to an identified or identifiable natural person or to data rendered anonymous in such a way that the data subject is not or no longer identifiable. This Regulation does therefore not concern the processing of such anonymous information, including for statistical and research purposes."

[6] National Human Genome Research Institute in NIH, Privacy in Genomics, April 21, 2015, available at https://www.genome.gov/27561246#al-5 (..."In 2013, as required by the passage of the Genetic Information Nondiscrimination Act, the Privacy Rule was modified to establish that genetic information is health information protected by the Privacy Rule to the extent that such information is individually identifiable, and that HIPAA covered entities may not use or disclose protected health information that is genetic information for underwriting purposes. There are no such restrictions on the use or disclosure of PHI that has been de-identified.")

[7] HIPAA プライヴァシー規則における匿名化の方法については、次の論文などを参照されたい。たとえば、佐藤智晶「米国と欧州における医療情報法制をめぐる議論」東京大学政策ビジョン研究センターワーキング・ペーパーPARI-WP, No. 9, Jan. 15, 2013, available at http://pari.u-tokyo.ac.jp/policy/working_paper/WP130115_satoc.pdf

っていたのに対し、2000 年の最終規則では次のように変更された。 第 1 に、匿名化された医療情報について、個人識別に用いられるという合理的な理由がない場合、と規定された（Health information that does not identify an individual and with respect to which there is no reasonable basis to believe that the information can be used to identify an individual is not individually identifiable health information）。すなわち、2000 年の規則では、結果責任でも厳格責任でもなく、合理性の基準で匿名化されているかどうかが判断されることになった。 第 2 に、個人識別可能な医療情報の匿名化として、次の 2 つの方法が明記された。当然ながら、上記 2 つの匿名化の方法は、ゲノムデータにも適用される[8]。

1 つめは、情報の受領者が個人を識別してしまうリスクについて、最小化されていることを専門家が確認する方法である。専門家は、統計的または科学的手法によってリス クが最小化されていることを確認するものと規定されている。このような柔軟かつより合理的な別の匿名化を許容する規定は、規則（案）にはまったく含まれていなかった。

2 つめは、18 種類からなる所定の情報を予め除去し、残りの情報では個人識別できないことを確認する方法である（いわゆる、セーフハーバー・ルール）。18 のデータとは、名前、州以下の住所、誕生日等の年月日、電話番号、FAX 番号、E メールアドレス、社会保障番号（SSN）、診療録番号、医療保険の受益者番号、銀行口座の番号、資格等の番号、自動車登録等の番号、医療機器番号、ウェブの URL、IP アドレス、指紋や声紋等の生体認証記録、顔面写真等のイメージ、その他の個人識別コードである。

なお、米国では一塩基多型が 30 から 80 あると個人識別性を帯びうるという指摘があり[9]、逆にいえば、30 未満であれば個人識別性がないゲノムデータに該当しうることになる。個人識別性のないゲノムデータであれば、当然ながら HIPAA プライヴァシー規則の保護対象にはならない。

---

[8] See, e.g., Simson L. Garfinkel, De-Identification of Personal Information, National Institute of Standards and Technology Internal Report 8053, October 2015, available at http://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf

[9] たとえば、El Emam K. Methods for the de-identification of electronic health records for genomic research. Genome Medicine. 2011;3(4):25. doi:10.1186/gm239, available at http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3129641/ (“There is evidence that a sequence of 30 to 80 independent single nucleotide polymorphisms (SNPs) could uniquely identify a single person”). See also Lin Z, Owen A, Altman R. Genomic research and human subject privacy. Science. 2004;305:183. doi: 10.1126/science.1095019.