

How Anonymous is Anonymous Enough? Lessons on Longitudinal Health Data from IQVIA's €7 Million Fine for Data Privacy Violations

October 14, 2026

By: Alfred Brunetti

Any organization that treats “de-identified” or “anonymized” health datasets as outside the scope of data privacy laws should revisit that assumption now.

Introduction

On October 2, 2026, the Italian Data Protection Authority (the *Garante per la Protezione dei Dati Personali*, or *Garante*) announced a decision¹ and accompanying €7 million fine against IQVIA Solutions Italy S.r.l., a wholly-owned subsidiary of IQVIA, the well-known global advanced health information provider. The *Garante* concluded that the health information of approximately one million patients, housed within an IQVIA database, was not anonymous (as IQVIA had argued) but, rather, it was personal health data subject to the full weight of the General Data Protection Regulation (GDPR). While this particular decision may only be binding in Italy and is not controlling authority for U.S.-only companies, the inquiries made by the *Garante* are the same critical questions that HIPAA regulators, the Federal Trade Commission (FTC), and state attorneys general are increasingly asking of life sciences companies, health data analytics firms, and their vendors: Can data that remains tied to a personal identifier actually be treated as de-identified? And who is accountable for personal data as it flows from collection to commercial users?

The evaluation of de-identification rigor and examination of dynamic responsibilities across a data supply chain at the center of this recent *Garante* decision have become a primary focus for U.S. data privacy authorities.

What Happened?

The *Garante* investigation followed inspections conducted in April 2025 and a data breach notification and focused largely on Longitudinal Patient Data (LPD), which comprises information collected through repeated observations of the same patient over an extended period of time. IQVIA — the self-proclaimed “world leader in clinical research and healthcare data” — maintains a LPD database containing records derived from general practitioners' electronic medical record systems and used to support studies commissioned by pharmaceutical companies. Each patient record was associated with a persistent “Pat ID,” which meant the same individual could be followed across visits and years.

IQVIA argued that the database contained only anonymous information and therefore did not contain personal data and, thus, could not be subject to GDPR regulation. The *Garante* rejected that position and found that combining a stable patient-level identifier with granular clinical and demographic detail (like year of birth, sex, diagnoses, symptoms, prescriptions, laboratory tests, vaccinations, and location data) allowed individual patients to be singled out and, by

reasonable means, re-identified. Records reached back to 2001, with no defined retention or deletion schedule. Separately, the *Garante's* analysis of an IQVIA self-reported breach found that IQVIA had exposed the identifying information (including names and contact details) of roughly 3,370 patients, the vast majority of which also had health data in the affected set.

The *Garante* also held that, from the moment data was first collected from the providing physicians, IQVIA acted as a controller, not a mere processor or downstream recipient. This determination was based on several factors, including ownership and use of the database, the structure of the contractual arrangements, IQVIA's role in instructing the software provider that performed the anonymization “add-on,” and the fact that the processing served IQVIA's own commercial purposes. From that determination, the Authority found violations of GDPR Articles 5, 9, 13, 25, 28, 32, and 35, spanning lawfulness, storage limitation, integrity and confidentiality, accountability, special-category data, transparency, privacy by design, processor engagement, security, and the failure to conduct a data protection impact assessment (DPIA). The decision gives IQVIA approximately four months to document how it will bring its program into compliance or to restructure it entirely. The €7 million penalty was driven by the volume and sensitivity of the data, the duration of the conduct, and IQVIA's level of cooperation.

What Do We Need to Know?

Pseudonymization is not anonymization.

The analytical core of the decision is the *Garante's* position that replacing a name with a code does not make data anonymous where re-identification remains reasonably possible and, thus, pseudonymized data remains personal data. The *Garante* faulted IQVIA for relying passively on generic assurances about the anonymization process rather than performing its own proactive and retrospective testing of numerous elements including, for example, the software itself, free-text clinical fields that can reintroduce identifiers, the results of rare attribute combinations, and the heightened exposure risks that accompany large-scale processing of sensitive data. This *Garante* decision highlights and distinguishes an important nuance of the Court of Justice of the European Union's landmark September 2025 decision, *European Data Protection Supervisor v. Single Resolution Board* (C-413/23 P). That Court decision held that pseudonymized data is not automatically personal data for *all* parties but, rather, that such classification depends largely upon whether the recipient can reasonably re-identify the data to an individual. The Court went on to explain that the central elements to be considered for such a classification examination are the technical, organizational, and legal factors that impact the data. A confluence of these control factors, some deficient and some absent, helped inform the *Garante's* determinations.

U.S. practitioners will recognize the general structure of such a posture. Under HIPAA, certain information is deemed “de-identified” only where it no longer identifies an individual, and there is no reasonable basis to believe it can be used to identify one. Critically, HIPAA permits a re-identification code only where the code is not derived from or related to information about the individual, cannot otherwise be translated to identify the individual, and is not disclosed or used for another purpose. Similarly, many state-specific comprehensive data privacy laws define the term “de-identified” as requiring the satisfaction of certain contractual, public-facing, and control-based obligations. A persistent longitudinal patient identifier enriched by clinical detail is exactly the fact pattern that strains both the expert-determination standard and the re-identification-code condition.

Controller status follows commercial purpose, not paperwork.

The finding from this decision that is likely to have the broadest operational impact is the *Garante's* determination that IQVIA was a controller from the point of collection. Often, companies that receive provider-derived data frequently characterize themselves as processors, service providers, or recipients of already-anonymized output. The *Garante* looked past labels and examined who designed the data flow, who directed the technical transformation, and whose business the processing served. U.S. law asks functionally similar questions. HIPAA's general rules expressly prohibit covered entities from using or disclosing protected health information except as permitted, and limit business associates to the uses

permitted by their contract or required by law. Similarly, a well-drafted data processing addendum — an increasingly fundamental part of many business contracts — will specify the permissible and prohibited uses for personal data that may flow during the business relationship.

The U.S. overlay for non-HIPAA entities.

Many life sciences and digital health companies, health data analytics firms, and data licensors are not regulated by HIPAA because they are neither a covered entity nor a defined business associate. Nevertheless, such entities may be subject to federal enforcement arising from their treatment of certain personal data. The FTC's Health Breach Notification Rule applies to vendors of personal health records (PHR), PHR-related entities, and their third-party service providers — all entities squarely outside of HIPAA's scope. It clarifies that a “breach of security” includes unauthorized disclosures, not just security incidents. The FTC's *GoodRx* and *Premom* enforcement actions remain the clearest U.S. precedents for health data shared with advertising and analytics partners contrary to a company's own representations, and the Commission paired those theories with Section 5 of the FTC Act, 15 U.S.C. § 45(a), which prohibits unfair or deceptive acts or practices.

A company that has convinced itself its dataset is anonymous has, in effect, pre-decided the re-identification question that drives that analysis — and the IQVIA matter illustrates how badly that can go when a regulator disagrees.

Guidance from the New Jersey Division of Consumer Affairs on protecting patient information reinforces the same themes the *Garante* emphasized: minimize collection, exercise caution with website tracking and marketing tools that share user data with third parties, conduct independent cybersecurity audits, encrypt, control access, train employees, and adopt retention policies that cull unnecessary sensitive data.

What Do We Need to Do?

Companies that build, license, or rely on provider-derived or patient-level health datasets should treat the following as near-term priorities:

1. **Reassess** every dataset your organization classifies as “anonymized” or “de-identified,” starting with patient-level data. Legal, privacy, and data science leads should jointly document the basis for each classification and examine whether a persistent patient, member, or device identifier lets records be linked over time.
2. **Determine and document** your role in each data flow (e.g., sole or joint controller, processor, service provider, covered entity, etc.), and assume a regulator will judge that role by who directs the processing and who benefits from it. Then align your contracts with those roles.
3. **Conduct** or refresh data protection assessments for large-scale health data processing, and build in purpose-specific retention and deletion schedules. Remain mindful of the fact that the *Garante* cited as separate violations both a missing DPIA and the open-ended retention of records dating back to 2001.
4. **Confirm** the legal basis and transparency chain for provider-sourced data. Where data originates with treating providers, document in writing who is responsible for informing patients and obtaining any required consent, whether under HIPAA or more comprehensive sensitive data legal regimes.

The *Garante's* decision is the latest wake-up call that the “it's anonymized” answer is no longer accepted at face value by increasingly sophisticated and detail-oriented regulators. As long as data identifiability remains at the center of investigative and enforcement focus, companies whose business models depend on data classifications should be prepared to defend their determinations with documentation, testing, and contracts rather than assumptions.

¹Provision No. 710 of September 23, 2026.