MDDX
RESEARCH & INFORMATICS

WHITE PAPER

RECORD RETENTION: Imaging Data Longevity

MDDX RESEARCH & INFORMATICS
580 California St, Floor 16
San Francisco, California 94104

T (800) 441-MDDX
F (866) 382-4696
info@mddx.com
www.mddx.com

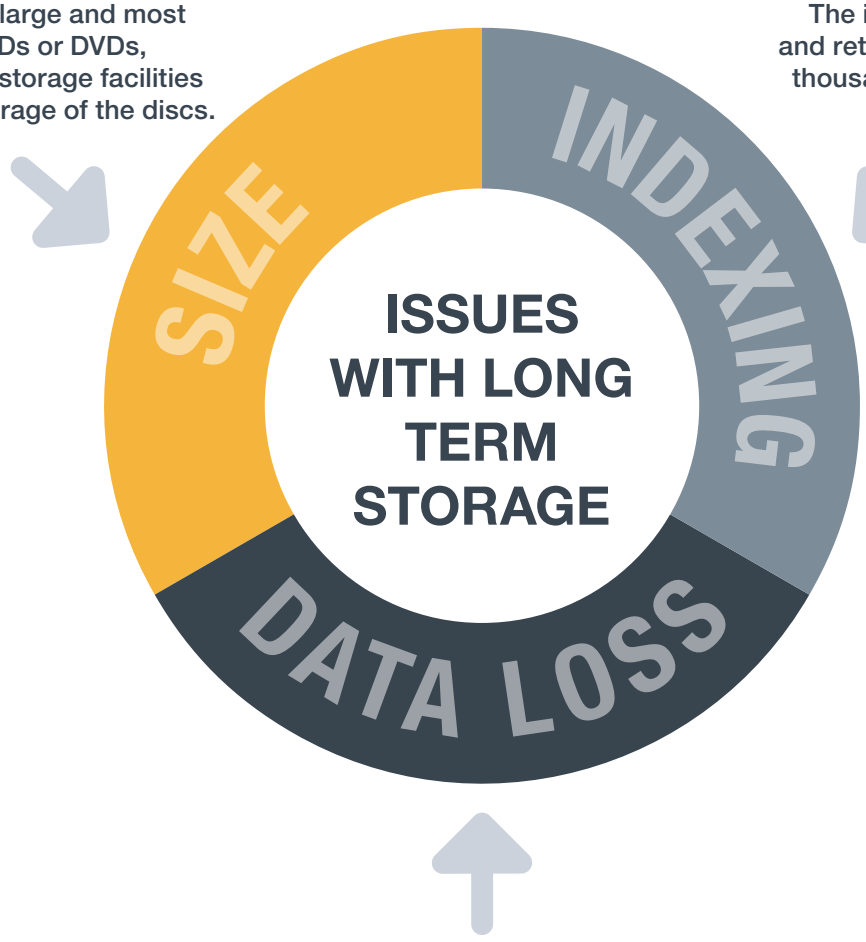
RECORD RETENTION: IMAGING DATA LONGEVITY

A common line of questions posed in vendor audits focus on the record retention policy for records generated during a clinical trial. Most research group policies cover standard data, validation records, and other text-based data. However, when it comes to the retention of medical images acquired during the trial, the long-term storage becomes more challenging and therefore overlooked, and often not in concordance with the group's own stated record retention policy.

HISTORICALLY, LONG-TERM STORAGE OF CLINICAL TRIAL IMAGE DATA HAS BEEN HAMPERED BY SEVERAL DISTINCT FACTORS:

Image files are large and most were sent on CDs or DVDs, requiring large storage facilities for physical storage of the discs.

The indexing, management, and retrieval of data stored on thousands of discs was a job of a daunting task.



When stored on hard drives, the potential for damage or corruption increases as the hard drives age. Finally, incomplete or improper de-identification of Protected Health Information (PHI) on images greatly increases the potential for unauthorized use or disclosure. With the increase in regulations, fines, monitoring, and enforcement, this creates a significant liability for research groups seeking to comply with retention periods.

WHY RETAIN DATA

The two key reasons for retaining clinical trial data are: use (current and future), and to satisfy health authority regulations. In larger institutions and companies, there are usually dedicated compliance officers monitoring current legislation and best practices to ensure that their organizations are satisfying the regulations. However, in small to mid-sized companies as well as academic/non-profit research institutions, the steadfast adherence to changing policy is not as commonly emphasized.

While achieving regulatory compliance is the primary reason most cited for data retention, a more meaningful purpose also exists: Images which are indexed and then stored properly, protecting data integrity and privacy, become a lucrative asset fueling future discoveries, confirmatory research and novel inventions. This is especially true for organizations actively performing research and development: the retention of records builds an inventory of collective knowledge creating longevity to the research being invested in today. With many discoveries resulting as a side-effect or offshoot of another main area of research, creating this virtual knowledge vault enables rapid data harvesting, meta-analyses and data licensing opportunities.

> **"We went through litigation that required us to retrieve data from 1986 through 1988. It was quite helpful for us to have a system where we had ready access to the data."**
>
> *Senior Executive at Bayer*

RECORD RETENTION: IMAGING DATA LONGEVITY

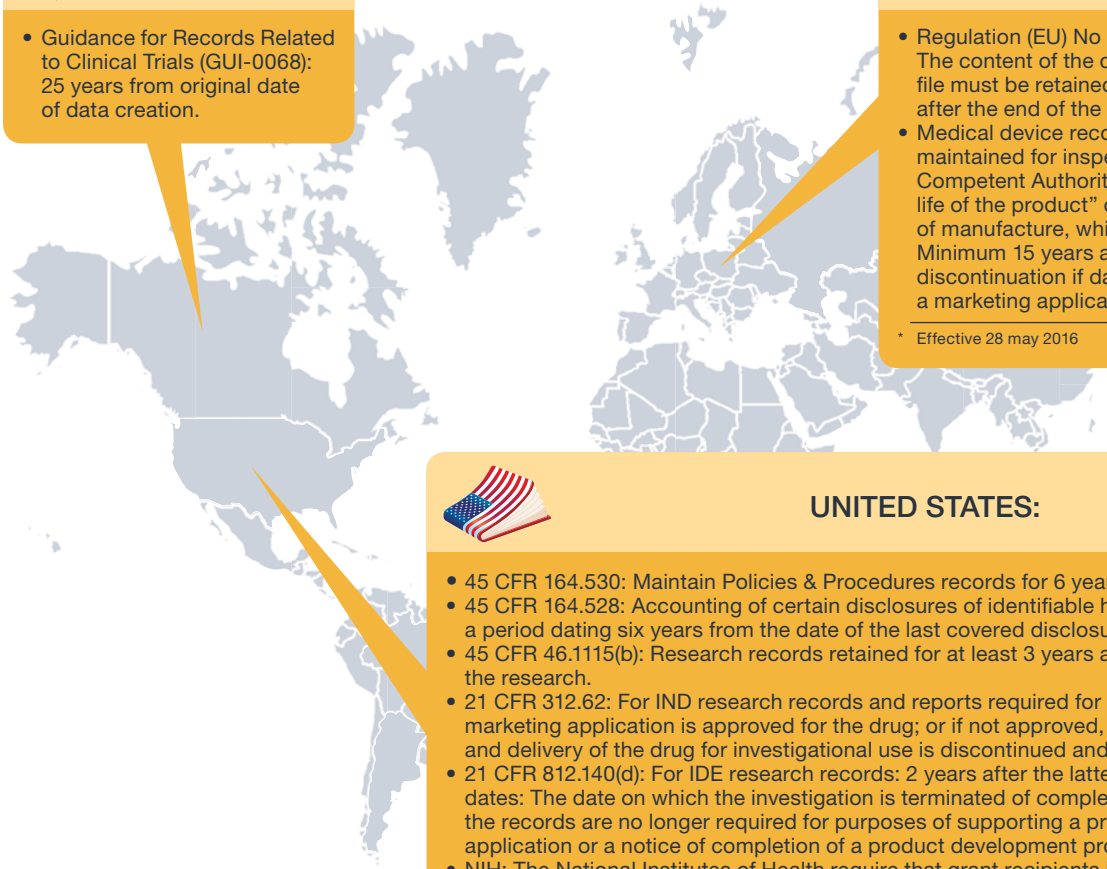
REGULATIONS

Record retention requirements under the law vary widely. Even within a single country, there exists a great deal of variability in the rules. The factors that determine the germane minimal retention period are the type of trial, funding origins of the trial, individual institution's policies, and sponsor's policies, the regulations can even vary within regulatory authorities themselves. One of the reasons for the variability is that the governance of clinical trial data falls into many of the following categories:

business records, policies and procedures, disclosure records, research records, medical device records, and the Trial Master File.

It is not surprising that the current regulations create confusion in determining the right retention time, as they range from: 2 to 25 years. Further confusing the issue, the retention clocks starts ticking at significantly differing time points: from date of creation to the end of the trial, to the date of the last disclosure.

CURRENT RETENTION PERIODS IN NORTH AMERICA AND EUROPE ARE DETAILED BELOW:



CANADA

- Guidance for Records Related to Clinical Trials (GUI-0068): 25 years from original date of data creation.

EUROPEAN UNION

- Regulation (EU) No 536/2014*, Article 58: The content of the clinical trial master file must be retained for at least 25 years after the end of the clinical trial.
- Medical device records must be maintained for inspection by the Competent Authorities for the "useful life of the product" or 5 years from date of manufacture, whichever is greater. Minimum 15 years after Clinical Trial discontinuation if data is used to support a marketing application

* Effective 28 may 2016

UNITED STATES:

- 45 CFR 164.530: Maintain Policies & Procedures records for 6 years after last use.
- 45 CFR 164.528: Accounting of certain disclosures of identifiable health information for a period dating six years from the date of the last covered disclosure:
- 45 CFR 46.1115(b): Research records retained for at least 3 years after completion of the research.
- 21 CFR 312.62: For IND research records and reports required for 2 years after a marketing application is approved for the drug; or if not approved, 2 years after shipment and delivery of the drug for investigational use is discontinued and the FDA so notified.
- 21 CFR 812.140(d): For IDE research records: 2 years after the latter of the following two dates: The date on which the investigation is terminated of completed, or the date that the records are no longer required for purposes of supporting a premarket approval application or a notice of completion of a product development protocol.
- NIH: The National Institutes of Health require that grant recipients keep all data for three years beyond the time that the final expenditure report of the grant is reported.

With the wide international variability in record retention periods, the conservative approach is to implement a simplified record retention period equal to the maximum period of any one Health Authority regulation. By instituting a 25-year policy, the majority of organizations will safely be within the current regulatory requirements while also providing ample time for any secondary discoveries or data harvesting to occur.

CASE STUDY

In a recent project MDDX performed for a cardiac device manufacturer, our team was asked to collect and de-identify over 21,000 legacy DICOM datasets stored on CD/DVDs, hard drives, and servers at multiple imaging core labs. This imaging asset comprised the key image data supporting two landmark trials for this multi-billion dollar company. Our analysis discovered that approximately 4% of the images had been corrupted during storage, and were unusable. More troubling was the discovery that only 19 of the DICOM datasets (<0.01%) were properly de-identified. Therefore, rather than creating an easy-to-access resource of medical images, both the sponsor and the core labs were sitting on a large potential liability of hard to retrieve images locked in back closets. MDDX reduced their PHI risk and increased the value of their data asset by retrospectively de-identifying all viable DICOM and indexing them in our secured Image Repository – which provides immediate online access for authorized sponsor staff. Their locked away liability was transformed into an easily accessible and robust asset.

IMAGE REPOSITORIES

Researchers who have conducted clinical trials involving imaging endpoints understand the challenges that come with collecting high-quality medical image data. Variations in acquisition techniques, non-adherence to the protocol, non-compliance to DICOM standards, improper DICOM handling and unavoidable image quality issues combine to make imaging a complicated and expensive endpoint for any trial.

Building a long-term de-identified image repository has many benefits. At a minimum, the repository allows sponsors to satisfy all health authority regulations regarding record retention. For forward-thinking companies, the repository has expansive potential to elicit a multitude of benefits including:

- Range finding of imaging end points for future trials
- Determining performance metrics of your research sites for future trial selection or risk-based monitoring
- Use of the images for marketing or for training clinical specialists and physicians
- Use of the images in aiding the development of software-based visualization and/or diagnosis of the condition
- Initiating new sub-studies or exploratory investigations not in the original scope (these sub-studies could commence next year or 10 years from now)

MEDICAL IMAGES AS SOURCE DATA

The core elements of clinical trial regulations are record keeping and audit trails. Therefore the growing best practice is for all data to be retained. All records created, generated, or referenced can and should be retained. These records range from policies and procedures, to staff records, validation records and all clinical activities. The majority of this information is defined as either source data or electronic source data.

Table 1: Definition of source data

Source data	All information in original records and certified copies of original records of clinical findings, observations, or other activities in a clinical investigation used for reconstructing and evaluating the investigation.
Electronic source data	Data initially recorded in electronic format. They can include information in original records and certified copies of original records of clinical findings, observations, or other activities captured prior to or during a clinical investigation used for reconstructing and evaluating the investigation.

Medical images in the DICOM format are inherently electronic source data. Due to the fact that medical images can be accidently truncated, corrupted, over de-identified or intentionally manipulated, the ability to trace the authenticity of a medical image to its source is a critical factor in record retention. Additionally, medical images must carry a measure of data integrity as they moves from research site to imaging core lab and finally to the image archive.

SUMMARY AND RECOMMENDATIONS

The combination of industry trends, regulatory requirements, good laboratory practices, and prospective business possibilities point firmly towards creating and implementing long-term record retention solutions. Several universities and institutions have begun leading the charge, encouraging both long-term storage and de-identification, recognizing the opportunities for future research and findings between departments and even between institutions. This two-pronged approach of storage and de-identification meets the best practice standards for international trials while also laying the groundwork for future breakthroughs, discoveries, and growth.