Long-Term Archiving Guidance for Analytical Instrument Data

Version 3.0

March 31, 2025

Public Interest Incorporated Association

Japan Image and Information Management Association

R&D Data Archiving Committee

Contents

1. Objectives

The purpose of this guidance is to present a method for transferring electronic data (hereinafter referred to as 'analytical instrument data') output from various analytical instrument at research facilities, laboratories, etc. in the pharmaceutical industry from the instruments to storage locations while maintaining their reliability and for ensuring their long-term and secure archiving and management, assuming that the data may be re-processed.

2. Scope of Application

This guidance covers analytical instrument data handled in the pharmaceutical industry for the following purposes:

a. Data used for submission and reporting to regulatory authorities
b. Data required to be archived under GxP regulations[1]
c. Data thought to need to be archived at the facilities

It is assumed that the analytical instrument data may be re-processed. If data output from analytical instruments such as a balance, pH meter is determined in its as is and not re-processed, it is not included in the subject of this guidance.

3. Introduction

With rapid changes in the IT environment, the number of electronic records that must be archived is rapidly increasing. On the other hand, unlike paper-printed records, electronic records are feared to become unreadable in the future, such as through migrating to a new system and software upgrade.

In the case of outsourced studies, when transferring analytical instrument data between laboratories / facilities, paper-printed data can be easily transferred. But in the case of analytical instrument data itself, as there are no clear standards on how to deliver the data with authenticity[2] and how to archive the transferred analytical instrument data, the facilities are forced to respond in an optimal procedure that they considered themselves.

In the pharmaceutical industry, this problem is more serious because analytical instrument data must be archived for a long period of time in a state where the data is reliable and can be re-processed (see Section 3.3).

The premises and background that led to the development of this guidance are provided below.

3.1. Overview of Analytical Instrument Data

Analytical data and processing results as well as analytical conditions, processing conditions, sample sequences, electronic signatures, audit trails and logs of the system are stored in the analytical instrument or the

---

[1] Any of the standards established by the regulatory authority to ensure the safety of patients and the reliability of the studies. Representative laws of the pharmaceutical industry are exemplified, but not limited to:
GLP: Good Laboratory Practice
GCP: Good Clinical Practice
GMP: Good Manufacturing Practice
GVP: Good Vigilance Practice

[2] Reference: Pharmaceutical and Food Safety Bureau Notification No. 0401022 dated April 1, 2005: Use of Electromagnetic Records and Electronic Signatures in Applications for Drug Approval or Licensing, etc. (ERES Guidelines), 3.1.1. Authenticity of Electromagnetic Records.

computer connected to the analytical instrument. Recently, it is often stored in a data storage server, a data management system, or the like via a network.

In this guidance, the following terms are used for the analytical instrument data.

There are two storage formats of analytical instrument data, A., B. below, and the analytical instrument data includes three types of data, a., b., and c. below.

A. Original data

Electronic data output from the analytical instrument. It is also called the original record and is part of raw data under GxP regulations. It is often output in an instrument-specific format. Depending on the instrument, it may or may not include derived data and metadata other than analytical data.

B. Standard format data

Analytical data, some derived and metadata which are converted to a standard format such as AIA. The remaining derived data and metadata are missing without being converted, but by storing the missing data together with materials to supplement the missing data, they can be considered equivalent to the data before conversion.

a. Analytical data

Data obtained from analytical instruments. Chromatogram data such as high-performance liquid chromatography (HPLC), spectrum data such as infrared spectrophotometer (IR), nuclear magnetic resonance equipment (NMR) and mass spectrometer (MS), and weighing data of balance, etc.

b. Derived data

Integrated results (e.g. peak area) and calculated results (e.g., concentration), etc. using original data.

c. Metadata

Information of units etc. supplemented to clarify the meaning of the data, the date and time of data acquisition, information identifying analytical instruments, analytical conditions, audit trails, etc.. The FDA (Food and Drug Administration) specifies metadata in its data integrity guidance.[3] In this guidance, data related to analytical instrument data and audit trails are presented separately, referring to the FDA's idea. That is, analytical conditions, processing conditions, sample sequences, and the like are denoted as analytical metadata, and the audit trail is denoted as audit trail metadata.

3.2. Problems with Analytical Instrument Data Printed on Paper

Traditionally, in the Japanese pharmaceutical industry, it has been common practice to print and store analytical instrument data and processing reports on paper. However, to solve the following problems and ensure data integrity, it has recently been required to store electronic data.

a. The printed data cannot be re-processed.

b. All data that are not normally printed (processing conditions, analytical conditions, etc.) and data that cannot be printed (three-dimensional data, etc.) are lost, even if they are included in the electronic data.

[3] Source: Data Integrity and Compliance With Drug CGMP Questions and Answers Guidance for Industry (FDA, Dec. 2018).

c. Large space to store paper is required.

3.3. Concerns on Analytical Instrument Data

3.3.1. Re-Processing

It may be required in the pharmaceutical field to re-process the analytical instrument data several years later from the actual acquisition. One example is the case of re-processing of the related substances test. It may be confirmed how much the related substance focused on the specified lot were included in the past several lots. In this case, it may be necessary to re-process past analytical instrument data to investigate unreported micro peaks.

In this context, since there may be differences between the retention period of analytical instrument data and the frequency of software version upgrades for the instruments, issues may arise when re-analyzing data that has been preserved over the long term. In addition, for various reasons it may sometimes be necessary to switch to analytical instruments from other manufacturers. (Figure 1)

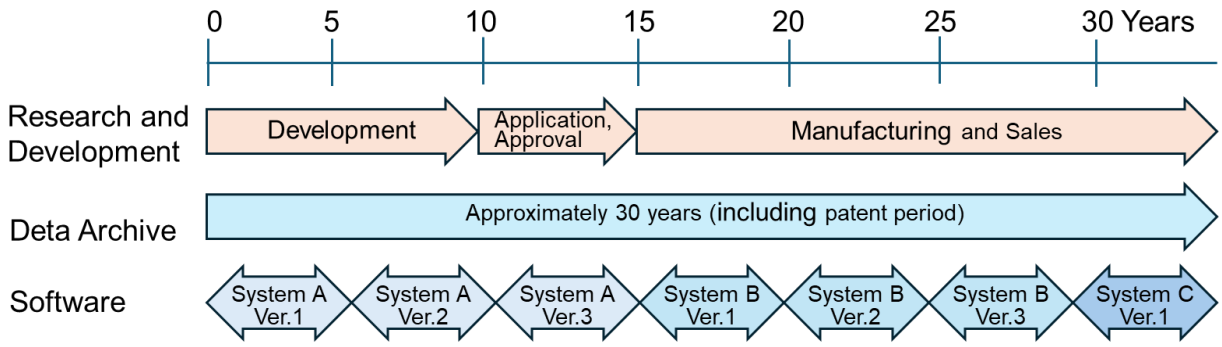

Figure 1. Relationship Between the Data Archiving Period and Software Upgrade[4]

When analytical instrument data is transferred to instrument software from another manufacturer, identical results cannot be guaranteed due to differences in the software algorithms.

Further, even if the software of the same manufacturer is used, when the version is different the same result as the original analysis result is not always obtained.

The reasons are as follows.

a. The algorithm has changed. (Including improvement of calculation precision)
b. The computer is using the binary (not the decimal) and a limited digit number for the calculation. So, the last digit contains an error. The calculation results also contain errors.
c. The computer (and operating system) is now moving from 32-bit to 64-bit. It may cause a difference in the results.

To avoid these situations, it is necessary to store not only the analysis software but also the computer suite,

[4] Source: "Creating Electronic Data Storage Packages and Assurance of Their Reliability," JIIMA Digital Document 2024 Webinar, November 2024.

including the operating system, for the re-processing environment. But, this is not realistic.

3.3.2. Data Migration and Long-Term Archiving

Considering the need for re-processing, it is desirable for the pharmaceutical industry to store analytical instrument data in a state that it can be re-processed for 30 years.[5] However, if hardware updates or software upgrades are performed, the processing software may not be able to accommodate a new operating system (OS) and analytical instrument data stored over a long period of time may not be available. In addition, when support for the database software being used ends, it may not be possible to migrate the analytical instrument data to the system of another manufacturer.

Typical methods for the long-term archiving of analytical instrument data are illustrated in Figure 2.

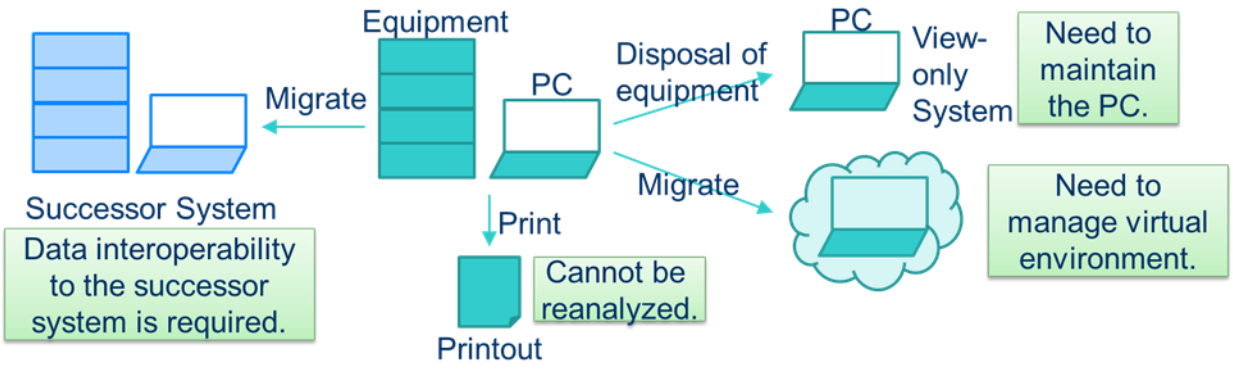

Figure 2. Long-Term Archiving Methods for Analytical Instrument Data[6]

It should be noted that printed copies of analytical instrument data cannot be reused (see lower center of Figure 2) and are therefore excluded from the scope of this guidance. Rather, this document addresses concerns regarding other possible countermeasures.

A) Physical storage of the processing computer (see upper right of Figure 2)

The relevant concerns are listed below:

a. Increase in managed systems
b. Mechanical lifetime (computer body, printer, external storage)
c. Maintenance expenses (including licensing expenses)
d. Storage place
e. Periodic check
f. Maintain backup of electronic data

[5] Example: Ministerial Ordinance on Standards for Manufacturing Control and Quality Control of Drugs and Quasi-drugs (Ministry of Health, Labour and Welfare Ordinance No. 179 of 2004) (GMP Ministerial Ordinance) Article 20, item 3, Article 30, etc.

[6] Source: "Data Storage Methods and Examples under Japan's GLP Regulations," 2024 KSQA International Conference, dated October 24, 2024 (https://jsqa.com/kenkyu/society/2024ksqa/).

g. Securing operators who can operate old systems

B) Virtualize and store the processing computers (see lower right of Figure 2)

The relevant concerns are listed below:

a. Increase in managed systems

b. Maintenance expenses (License maintenance costs are required; however, as many products are now sold under subscription contracts, it is necessary to take into account how the license will be maintained after the contract expires.)

c. Periodic checks

d. Maintain backup of electronic data

e. Securing operators who can operate old systems

f. Measures to be taken when additional hardware is required, such as hardware keys[7], interface cards[8], etc.

g. Licensing form of the software subject to virtualization[9]

h. To be fully compatible

C) Migration from current to succeeding (or compatible) systems (at left of Figure 2)

The relevant concerns are listed below:

a. Restrictions on migratable systems
It is difficult to know the constraints when introducing the system.

b. Constraints on the migratable range
For example, audit trails are not migratable.

D) Data migration via standard formats. (at left of Figure 2)

The relevant concerns are listed below:

a. Regardless of whether automatic or manual, the ability to export data and metadata is essential for a migration source system, and the ability to import them is essential for a migration destination system.

b. Loss of links to metadata or difficulty of the migration of links.

c. It is difficult to keep data integrity due to differences in the data management level between the migration source system and the migration destination system. For example, no audit trail is recorded other than the data generated by the system.

In this guidance, we recommend "D) Data Migration via Standard Formats." The reasons for this recommendation are described in Chapter 4. In addition, views on methods for transferring measurement data that lack a standard format are also presented.

[7] Hardware (dongle) such as the USB device used for use control such as software to prevent fraudulent access and fraudulent copying.

[8] An extended card adding an input-output interface to a computer.

[9] The hardware rather than users may need licensing (e.g., Microsoft Windows DSP version, OEM version).

3.3.3. Data Integrity

On the premise of re-processing, it is important that the reliability should be maintained when the analytical instrument data is stored or the storage location is moved. Specifically, for the analytical instrument data used for application and reporting to the regulatory authority, the data integrity has come to be required. To ensure this data integrity, it is required to store not only analytical data but also derived data and metadata (analytical and audit trail) under security management while maintaining legibility, but there are difficulties in meeting the requirements, including the concerns of 3.3.1 and 3.3.2.

Excerpts of the guidance with regulatory requirements are provided below.

Regulatory Requirements for the Storage of Electronic Data

The FDA has published the following Q&A in the guidance of "Questions and Answers on Current Good Manufacturing Practices, Good Guidance Practices–Records and Reports."[10] The answers were extracted only from the relevant parts and arranged for easy understanding.

Q) How do the part 11 regulations and "predicate rule requirements" (in 21 CFR part 211) apply to the electronic records created by computerized laboratory systems and the associated printed chromatograms that are used in drug manufacturing and testing?

A) The printed chromatogram would not be considered an exact and complete copy of the electronic raw data used to create the chromatogram, as required by § 211.68. The chromatogram does not generally include, for example, the injection sequence, instrument method, integration method, or the audit trail, of which all were used to create the chromatogram or are associated with its validity. Therefore,

a. The printed chromatograms used in drug manufacturing and testing do not satisfy the predicate rule requirements in part 211.

b. The electronic records created by the computerized laboratory systems must be maintained under these requirements.

FDA guidance[3] requires ALCOA for data integrity, and PIC/S guidance[11] requires CCEA as well. It is necessary to establish the ways to satisfy the requirements of the ALCOA+ prepared by integrating both of them so that it can deal with regulatory inspections (e.g., document-based compliance review) in response to the expectations of the regulatory authorities in Japan, the United States, and Europe.

ALCOA

Attributable: Signing and sealing, etc. Identify the attribution and responsibility of the data.
Legible: Data is recorded concisely and clearly so that it can be easily read.
Contemporaneous: Data is recorded at the time the work is performed and without delay.

[10] Source: Data Integrity and Compliance With Drug CGMP Questions and Answers Guidance for Industry December 2018 (https://www.fda.gov/media/119267/download).

[11] Source: P IC/S Guidance Good Practices for Data Management and Integrity in Regulated GMP/GDP Environments (PIC/S, PI 041-1 1 Jul. 2021) https://picscheme.org/docview/4234.

Original: Records are maintained as the original data. Accurate: Data and records are accurate and objective.

CCEA

Complete: Records have no data unavailability and are complete. Consistent: Records (data) are reasonable and consistent. Enduring: Record retention is durable. Available: Data can be retrieved as needed.

MHRA guidance[12] defines "true copy" as follows in section 6.11.2:

a. *A true copy may be stored in a different electronic file format to the original record if required, but must retain the metadata and audit trails required to ensure that full meaning of the data are kept and its history can be reconstructed.*
b. *Original records and true copies must preserve the integrity of the record. True copies of original records may be retained in place of the original record (e.g. scan of paper record), if a documented system is in place to verify and record the integrity of the copy. Organisations should consider any risk associated with the destruction of original records.*
c. *It should be possible to create a true copy of electronic data, including relevant metadata, for the purpose of review, backup and archival. Accurate and complete copies for certification of the copy should include the meaning of the data (e.g. date formats, context, layout, electronic signatures and authorizations) and the full GxP audit trail. Consideration should be given to the dynamic functionality of a 'true copy' throughout the retention period (see 'archive').*

In addressing regulatory requirements, risk management plays a critical role.

The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) engages in activities to promote harmonization in the three fields of quality, safety, and efficacy. In 2006, it issued the Guideline on Quality Risk Management, which was revised and published in 2023.[13] Risk management is implemented through a process that involves risk assessment (hazard identification, risk analysis, and risk evaluation), risk control based on the assessment, and a review of the results. Risk management is also an essential element in ensuring data integrity. It should be noted that the Fundamental Concepts of Quality Risk Management <G0-2-170>[14] in the Supplementary Information of the 18th Edition of the Japanese Pharmacopoeia provides an overview of the aforementioned guideline, which may be useful as a reference.

The Data Integrity Project of the GMP Subcommittee within the Quality Committee of the Japan Pharmaceutical Manufacturers Association (JPMA) has emphasized that addressing data integrity based on risk

[12] Source: 'GXP' Data Integrity Guidance and Definitions (MHRA, Revision1 March 2018).
[13] Source: ICH Q9(a) Guidelines for Quality Risk Management (https://www.pmda.go.jp/int-activities/int-harmony/ich/0049.html).
[14] Source: Basic Concepts of Quality Risk Management <G0-2-170> in the Supplementary Information of the Japanese Pharmacopoeia 18th Edition https://www.mhlw.go.jp/content/11120000/000788362.pdf.

analysis is extremely important in establishing an effective and efficient quality assurance system. Accordingly, in 2018 it published the Data Integrity Compliance Assessment Tool on its website.[15]
Likewise, when preserving or transferring analytical instrument data, it is necessary to conduct a risk assessment and take appropriate measures, considering factors such as the criticality of the data and the degree of human involvement in the process.

3.3.4. Data Format from Analytical Instruments

Analytical instruments generate various types of data. It is difficult to transfer electronic data saved in the original format (instrument-specific format) of one manufacturer's analytical instrument directly to an instrument of another manufacturer.

Therefore, in some cases, industry groups and other organizations have considered standard formats that are independent of specific manufacturers. However, since such a high level of versatility has not been required as in the case of image or audio data, many types of data have not yet been examined. While some instruments can output data in text-based numerical formats with high versatility, such file formats are also easily subject to tampering. Accordingly, when storing such data, it is preferable to use packaging tools or similar measures as explained in Section 5.1.3. In this context, chromatographic data, which is frequently used in analysis, is taken as an example. As the standard format of chromatographic data, the following formats have been proposed, but each has merits and demerits. So, it has not been unified yet.

a. AIA (Analytical Instrument Association)
   Standard format for chromatographic data, such as HPLC, defined by the US Analytical Instrument Association-AIA. Different versions may be incompatible.
b. NetCDF (Network Common Data Form)
   The NetCDF supports a machine-independent format for representing scientific data, as developed by Unidata, part of the University Corporation for Atmospheric Research-UCAR Community Programs (UCP).
c. JCAMP (Joint Committee on Atomic and Molecular Physical Data)
   Standard format for spectral data such as NMR, IR, and MS taken over by the International Union of Pure and Applied Chemistry-IUPAC.
d. AnIML (Analytical Information Markup Language)
   XML standard formats, as defined by American Society for Testing Materials. XML format not yet finalized due to drawbacks that the data size becomes 20 times bigger than the original one and the writing speed is slow even though the data acquiring speed from analytical instrument is fast.[16]

It should be noted that none of the above formats include all analytical metadata.

[15] Source: https://www.jpma.or.jp/information/quality/index_di.html.
[16] The XML format may improve the communication speed by improving the algorithm, but the modification of the algorithm is difficult because it had been developed by using machine language.

Chromatogram data can be converted to standard formats, but analytical metadata such as analytical conditions and processing conditions have different formats for each analytical instrument of different manufacturers. Also, since the processing algorithm is different for each manufacturer, it is not possible to completely reproduce the data from other manufacturers even if the processing parameters can be transferred. However, in the case of quantitative analysis, equivalent quantitative results can be obtained with other manufacturers' instruments when the relative values against the reference standards are almost the same. An example of such quantitative analysis is concentration calculation using a calibration curve created using reference standards.

However, even when using data in a standardized format, differences in data content may exist among manufacturers. In the AIA format, the structure for storing raw data together with various metadata is defined. Nevertheless, variations may occur in the metadata stored by different manufacturers, or in the units used for the stored data and metadata. In some cases, portions of the required data fields may be left blank, or units for detectors and time may depend on the specific instrument model (see *Appendix 1: Example of AIA Format Data*).

3.3.4.1. Interoperability of Analytical Instrument Data

Even when a standard format has been established, it is necessary to verify whether data migration between manufacturers using that format is possible. To this end, with the cooperation of the manufacturers, we examined the interoperability of data by taking HPLC and mass spectrometry (MS), both widely used in research facilities and testing laboratories, as case studies.

For HPLC, it was demonstrated that data in the AIA (Analytical Instrument Association) format can be transferred from one instrument to another and re-analyzed. Accordingly, it was shown that HPLC data in the AIA format may be mutually utilized among different manufacturers (see Chapter 5).

In contrast, with respect to MS data, it became clear that no common format usable across major manufacturers could be identified, and that data output from MS instruments of different manufacturers could not be re-analyzed. As a result, none of the formats were found to be interoperable.[17]

3.3.5. Use of Cloud Services

Traditionally, electronic data has been stored and managed within each facility using file servers, or application software installed on servers secured by the facility or in data centers. However, in recent years, various providers have begun offering their own application software (such as document management systems) and infrastructure (such as AWS and Azure) as services via computer networks, including the Internet. Such services are referred to as cloud services, the environment in which they are hosted is called a cloud environment, and the supporting technologies are referred to as cloud technologies. As a result, an increasing number of facilities now store electronic data and electronic documents in application software provided within cloud

[17] Based on a user questionnaire survey conducted with the cooperation of the Japan Society of Quality Assurance and the PDA Japan Chapter, netCDF (Network Common Data Form) was selected as a candidate for the standard MS format. Using MS instruments and applications from four different manufacturers, we confirmed whether data output/import was possible, but found that interoperability was difficult.

environments by service providers. In addition, shared sites on cloud environments are sometimes used for data transfer between facilities. For this reason, this guidance also provides explanations regarding cloud services.

In general, cloud services refer to businesses that provide a variety of services via the Internet. In addition to specific applications or storage, cloud providers offer a wide range of services, including servers and large-scale systems. Since cloud providers themselves handle system construction, maintenance, and operation, users can access services without the burden of additional effort or cost.

Cloud providers deliver services across hardware (infrastructure), middleware, and software via the Internet. Depending on the scope of these services, cloud offerings are broadly classified into three categories: SaaS, PaaS, and IaaS.

Cloud service delivery models include private clouds, which are built exclusively for individual users, and public clouds, in which multiple users share the same configuration. Operational models include single-cloud, hybrid-cloud, and multi-cloud approaches. A hybrid cloud integrates public clouds, private clouds, and on-premises systems into a single environment, while a multi-cloud approach involves contracting with multiple providers and using their cloud services in combination.

3.3.5.1. Long-Term Archiving of Analytical Instrument Data in a Cloud Environment

This section considers the regulatory requirements and practical points for long-term archiving of analytical instrument data in cloud environments.

In Japan, storing records in cloud environments within Good Laboratory Practice (GLP) facilities had long been prohibited. However, following the publication of the Addendum[18] to OECD Document No. 17 in June 2023, staff of the Pharmaceuticals and Medical Devices Agency (PMDA) stated at the 28th GLP Training Seminar that “PMDA does not deny the use of cloud technologies for the storage of GLP records.”[19]

In this guidance, it is assumed that analytical instrument data may be stored long-term in cloud environments. The service provider is responsible for maintaining and managing the relevant software (including change control) and the stored data (including security measures, backup, and restoration). Therefore, before adopting a cloud service, users are advised to evaluate the provider on a risk-based basis, considering aspects such as its quality management system, information security and business continuity, and the services offered, and to select services that fit the intended use.

Since cloud services require technical controls—such as data encryption—to ensure confidentiality and authenticity in an open environment, more requirements need to be verified compared with a closed environment. For further details, please refer to the relevant technical guidebook[20].

---

[18] Source: OECD Series No. 17 Supplement 1 on Good Laboratory Practice (GLP) Principles and Compliance Monitoring: Application of GLP Principles to Computerized Systems
https://jsqa.com/seikabutsu/open/glp_bukai/oecd-glp17_supplement1/.

[19] Source: Materials from the 28th GLP Training Seminar (https://www.pmda.go.jp/files/000264651.pdf).

[20] Source: See Long-Term Archiving of Analytical Instrument Data – Technical Guidebook, JIIMA 2021.

4. Policy

This chapter provides the concept behind the recommended methods for long-term archiving of analytical instrument data described in Chapter 5.

Table 1 provides a summary of the response policies to the concerns listed in Section 3.3.

Table 1. Response Policy to Address Concerns (Summary)

	Concern	Response Policy
1	At the time of re-processing, the processing algorithms are different for each manufacturer, so it is not possible to completely reproduce the data of other manufacturers even if the processing parameters have been migrated. (3.3.1)	Show in Section 4.1
2	At the time of re-processing, even with the same manufacturer, the same results as the original processing results may not be obtained due to differences in the software versions. (3.3.1)	Show in Section 4.1
3	When migrating data for long-term archiving, the link between original data and metadata is lost or migration of the link is difficult. (3.3.2)	Show in Section 4.2
4	When migrating data for long-term archiving, there is a limitation on the data management level in the system to which the migration is made and maintaining data integrity is difficult. For example, no audit trails are recorded other than the data generated by the migration system. (3.3.2)	Show in Section 4.3

4.1. Approach to Processing Results

As described in 3.3.1, analytical instrument data does not necessarily produce the same results as the original processing results. Even if the original data can be stored completely, it is permissible to include subtle differences depending on the available computer and software at the time of processing. Since the difference is assumed to be small, it is believed that it does not substantially affect the purpose of the test in many cases.

Therefore, it is necessary not to raise the requirement level more than necessary by clarifying what and how much to seek at the time of data long-term archiving. Possible requirement levels include the following:

a. The re-processing results need to be consistent with the original processing results.
b. If re-processing is possible, differences in processing results due to algorithm changes and other reasons are tolerated.
c. It is good if it can be reviewed from the previous and another viewpoint. For example, it is necessary to enlarge the area where the presence of a related substance is of concern and confirm the peak.

4.2. Approach to Migration and Long-Term Archiving

This guidance envisages 10 to 30 years as a period in which analytical instrument data is expected to be archived in a re-processed state. For this reason, the following workflow is proposed in this guidance (Fig. 3, Table 2).

The relevant workflows are described in detail in the Technical Guidebook and Operational Guidebook.[21] As a result, it is possible that the analytical instrument data can be independent from the analytical instrument to promote fluidization, and a contract service that provides a processing environment can be realized (Reference 7).

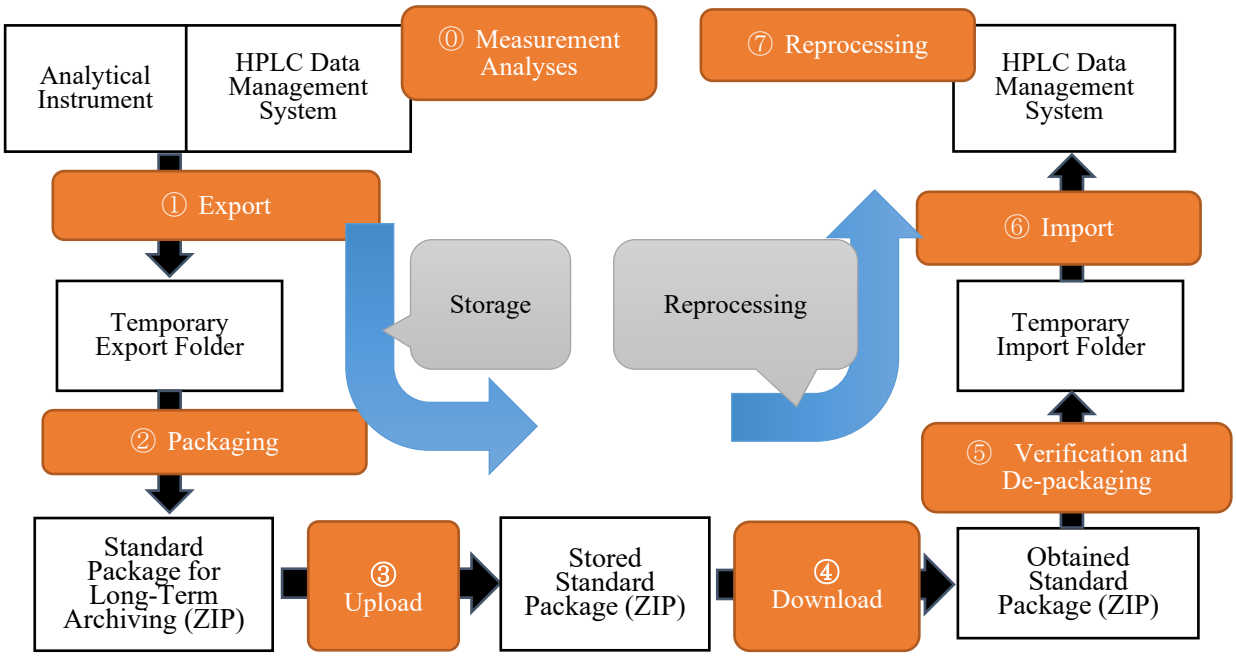

Figure 3. Workflow of Analytical Instrument Data

Table 2. Workflow from Analytical Instrument Data Acquisition to Reprocessing

Steps	Tools Used	Description	Purpose
⓪ Measurement and Analysis	Measurement Software Analysis Software	Measure and perform the first analysis.	
① Export	HPLC Data Management System	Output the data required for analysis to the export folder.	Archiving

[21] Source: See Long-term Archiving of Analytical Instrument Data – Operation Guidebook, JIIMA 2023.

Steps	Tools Used	Description	Purpose
② Packaging	Package Tools	Create metadata using files in the export folder and, if necessary, externally supplied information, and compile and record standard package files using ZIP. Consider setting and operation of access rights so that data is not tampered with from export to standard package creation. The export folder is a temporary area and is deleted after the package is created.	
③ Upload (Write)	Data Storage Server (WORM Media)	Upload the standard package file on the server. On the server side, record who was uploading. Consider setting and operation of access rights so that uploaded files are not tampered with. This may be stored in a WORM medium.	
④ Download (Read)	Data Storage Server (WORM Media)	Search and download the standard package of data to be reanalyzed. On the server side, record who downloaded and when. In some cases, data is read from a WORM medium.	Reprocessing
⑤Verification and De-packaging	Package Tools	Verify the standard package file, check the authenticity (falsification, etc.), and output the data in the standard package to the import folder.	
⑥ Import	HPLC Data Management System	Import and record data from the import folder. Consider setting and operation of access rights so that data is not tampered with from verification decompression to import. The import folder is a temporary area and is deleted after import.	
⑦ Reprocessing	Analysis Software	Reprocessing is performed.	

4.3. Approaches to Data Integrity

Data integrity is closely related to the original data, derived data, standard format data, analytical metadata, audit trail metadata, data processing processes, and OS/hardware (Fig. 4).

Exported and packaged analytical instrument data includes only analytical instrument data up to the point where the export and packaging task is performed.

In this guidance, it is recommended to export analytical instrument data on the condition that the third party confirms that data integrity is ensured under the specified conditions. The export and packaging tasks are performed at various times; for example, immediately after the completion of the analysis and processing or upgrading software.

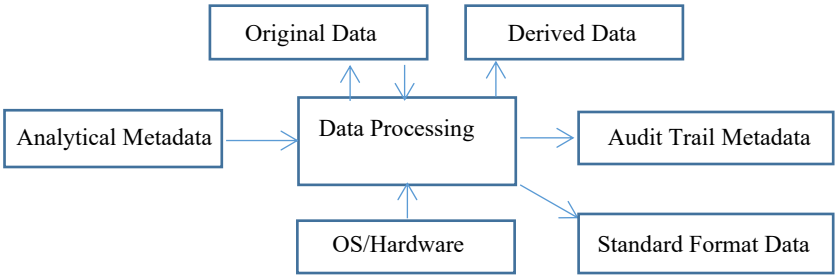

Figure 4. Elements Related to Data Integrity

After exporting analytical instrument data, data integrity is conditionally ensured. However, if the third party (QA, etc.) guarantees the appropriateness of the process from exporting the data to importing it into a new system, it is considered that the reliability of the data can be ensured even if the data export and import are not automated.

5. Recommended Migration and Long-Term Archiving Method for Analytical Instrument Data

This guidance provides a more detailed description of the process in which the standard format data from measuring instruments, as outlined in Section 3.1, is transferred to the data management system using the method described in Section 4.2, and subsequently reanalyzed in the HPLC data management system when necessary.

In this process, it is essential to establish mechanisms to prevent tampering with the data during the transfer from the measuring instruments that generated the data to the HPLC data management system where reanalysis is performed. The measure adopted for this purpose is data packaging.

5.1. Procedure for the Long-Term Archiving of Analytical Instrument Data

There are various types of analytical instrument data. In this guidance, chromatographic data was taken up as representative examples of analytical instrument data and the AIA format was taken up as its standard format, and methods for migration and long-term archiving of analytical instrument data were examined. In addition, consideration was also given to the handling of measurement instrument data for which no standard format exists.

5.1.1. Migration and Archiving AIA Format Data (HPLC Chromatograms)

Since it has been confirmed that AIA format data ensures interoperability between applications,[22] it is inferred that data reanalysis can be performed regardless of the application used. Therefore, this section explains the transfer and storage of data using the AIA format. It should be noted, as described in Section 3.1, that measurement instrument data includes measurement data, derived data, and metadata, and that all of these must be transferred and stored to ensure reproducibility.

If the AIA format output from HPLC were to include not only measurement data but also all derived data and

[22] It has been confirmed that HPLC systems from Nihon Waters K.K., Shimadzu Corporation, and Hitachi High-Tech Science Corporation can reproduce waveforms by reading data in the AIA format created by other companies and handle the data with their own analysis software. (Hitachi High-Tech Science Corporation's HPLC does not have a function to read the AIA format, so only writing was confirmed.)

metadata, then only the AIA format data would need to be transferred. In practice, however, this is not the case (an example of AIA format data is provided in *Appendix 1: Example of AIA Format Data*). Among the metadata, the sample schedule in particular is critical information that links the test items with the measurement results, but this cannot be associated with the AIA format data. Therefore, it is essential to store the AIA format data together with the derived data and metadata as a single package, in a tamper-proof manner.

In HPLC measurements, it is sometimes the case that samples prepared for purposes other than the primary objective are analyzed together with those prepared for the primary purpose. When storing a series of data under such circumstances, extracting only the data deemed necessary could raise suspicion of arbitrary selection. Therefore, the data should be stored on a measurement batch basis.

5.1.2. Migration and Archiving Analytical Instrument Data Without a Standard Format

For measuring instruments without an appropriate standard format, it is necessary to store the original format data in place of standard format data. Original format data is difficult to alter, contain metadata such as a creation date, and are useful for ensuring data integrity. Therefore, it is recommended to package and store the original data as is. In addition to the original format data, utility can be enhanced by also storing file output in text format (to make numerical values available for use) and files exported in PDF format (to ensure readability).

Unlike standard format data, original format data can only be analyzed with limited applications. However, as long as an application capable of analyzing the format is available, the original data can be used for reanalysis even after long-term archiving.

Furthermore, even if the software used at the time of measurement is no longer retained, reanalysis can still be supported by requesting assistance from the manufacturer, provided that the original data has been preserved. It should be noted, however, that given today's frequent corporate mergers among manufacturers, the availability of such support may pose challenges.

5.1.3. Analytical Instrument Data Package

An image of the analytical instrument data package is shown in Fig. 5 and 6.

The analytical instrument data package includes original data and standard format (AIA) data that can be archived for a long time while maintaining the dynamic state of the analytical data. By adding derived data (analysis results, concentration calculation results, files output in a human-readable format from the original data, analysis reports, etc.), analytical metadata, audit trail metadata, etc., the authenticity of the analytical instrument data is ensured.

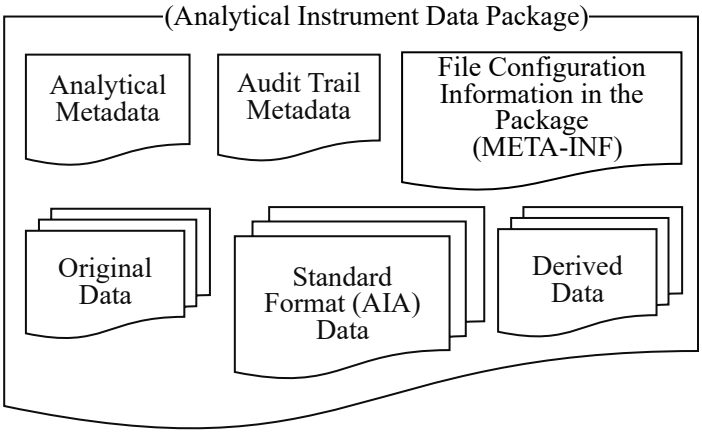

Figure 5. Packaging of Analytical Instrument data

It is recommended that the ZIP package be used as a package for these multiple files.[23] The ZIP package stores a special directory (META-INF) describing file configuration information in the package in addition to original data, AIA data, derived data, and metadata. This META-INF will be identified with a directory.

The specifications of the ZIP package are defined and presented in the Technical Guidebook.

The META-INF can also be managed outside the packages. In this case, the link is made with the hash value of the package. Appropriate techniques such as time stamps or electronic signatures are used to generate hash value. This is explained in the Technical Guidebook.

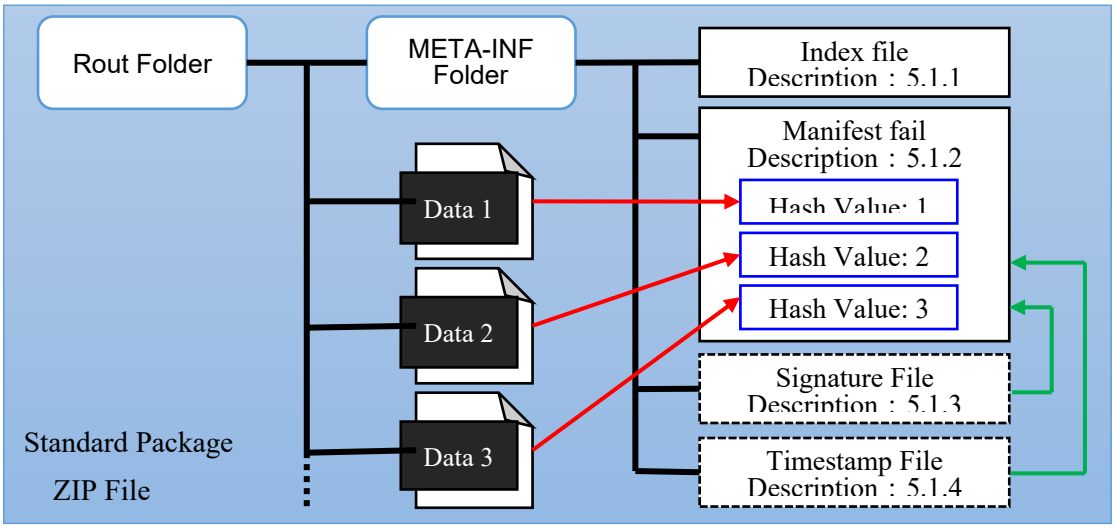

Figure 6. Examples of Structures in the Standard Package (See the Technical Guidebook)

[23] Other package candidates include ZIP-based purpose-specific packages, PDFs, and XML structures. Various ZIP-based purpose-specific package formats already exist (e.g., OPC for Office documents, EPUB for electronic books, and ASiC (Associated Signature Container, ETSI TS 102 918), which is being standardized in Europe). These formats can be considered to differ only in the definition of META-INF within the package. Defining a META-INF file specifically for measurement instrument data will also be a viable option in the future (in which case a new corresponding extension would need to be registered).

When a time stamp is used, it is necessary to establish a method for assuring effectiveness after the expiration of the certificate of the time stamp station. One method is to publish in newspapers, journals, etc., and use the National Diet Library to ensure readability.

In the case of re-processing, instead of directly manipulating the packaged data, the data is copied, and re-processing is performed using this copy, which makes it possible to eliminate the risk of erroneous erasure of original data or erroneous overwriting.

5.2. Ensuring Reliability of the Long-Term Archiving Process for Analytical Instrument Data

In the long-term archiving process of measurement data, some tasks are executed by systems (e.g., analytical instruments, tools, electronic document management systems), while others are performed by operators. For example, (i) in data export, the export itself is executed by the system, whereas the selection of data is performed by the operator; and (ii) in packaging, the creation of the standard package is carried out by the packaging tool, whereas the compilation of data and the creation of the hierarchical structure are performed by the operator.

Processes executed by systems are ensured through computerized system validation and related measures. In contrast, processes involving human intervention are subject to risks such as errors in data selection, operational mistakes in system handling, and even tampering with measurement instrument data. To mitigate the risk of such errors or tampering, it is necessary to ensure assurance through the preparation of work records and verification by a third party. Furthermore, in order to clarify the required records, items to be documented, and check points, it is desirable to prepare in advance planning documents or templates that specify the data subject to preservation, retention periods, and other relevant details.

For long-term archiving of data, it is also desirable to develop standard operating procedures (SOPs), preservation operation records, and third-party (e.g., QA) verification records, taking into account the operational practices of each organization. In addition, qualification and suitability checks of the storage media (e.g., physical media, servers) used for data preservation are also necessary. For further details, refer to Operational Guidebook 21.

Table 3. Method to Ensure the Reliability of Processes

Step	System Performance Assurance	Examples of Work Performed by Operators
⓪ Measurement and Analysis		
① Export	Validate Analytical instrument Check Exportable Data	Check Export Information Select Data to Export Execute Export
② Package Creation	Validate Tool	Select Data to Export Execute Export
③ Upload (Write)	Validate Upload Check Upload Location	Select Standard Package Select Upload Location

		Execute Upload
④ Download (Read)	Validate Download	Select Standard Package Select Download Location Execute Download
⑤ Verifucation and Unzip	Validate Verification and Unzip	Execute Unzip Validate Data, Check for Tampering
⑥ Import	Validate Analytical Instrument	Select Instrument to Import Select Data to Import Execute Import
⑦ Reprocessing		

5.3. Examples of Analytical Instrument Data to be Stored in the Package

The requirements for data storage vary depending on the purposes of use, the status of use, the level of reliability required, and the retention period. The purposes of use include reproducibility of results, compliance with document-based compliance review, and compliance with the principles of GxP.

Examples of measurement instrument data to be included in the package, depending on the purpose, are presented separately for cases where an appropriate standard format exists (Section 5.3.1, Table 4) and where no standard format exists (Section 5.3.2, Table 5). (For a detailed explanation of the purposes, refer to *Appendix 2: Explanation on the Purposes of Using Measurement Instrument Data*.)

5.3.1. Cases Where a Suitable Standard Format Exists

An example of analytical instrument data to be included in the package, in cases where an appropriate interoperable standard format such as the AIA format exists, is shown in Table 4. (For detailed explanations, refer to *Appendix 2: Explanation on the Purposes of Using Measurement Instrument Data*.)

Table 4. Example: Selection of Analytical Instrument Data for Each Purpose of Use (Standard Format)

Analytical Instrument data		Re-processing Only	Document-Based Compliance Review	GxP Regulations Compliance
Original Data (Instrument-Specific Format Data)		○	○	○
Standard Format (AIA) Data		○	○	○
Derived Data	Results (e.g., peak area) processed using original data etc.	－	○	○
	Calculated results (concentration) etc.	－	○	○
	Reports including analytical conditions	○	○	○
Analytical Metadata	Sample sequence	○	○	○

	Instrument parameters	○	○	○
	Processing parameters	○	○	○
Audit Trail Metadata	Original data-related	－	○	○
	Processing results-related	－	－	○
	System-related	－	－	－

○: Storage　－: No Storage Required

5.3.2. Cases Where a Suitable Standard Format Does Not Exist

An example of measurement instrument data to be included in the package in cases where no appropriate standard format exists is shown in Table 5. It should be noted that, for the purpose of reusing or reviewing measurement results at the time of measurement, utility can be enhanced by storing, together with the data, files output as numerical data in text format in place of AIA format data, as well as files exported in general-purpose formats such as PDF.

Table 5. Example: Analytical Instrument Data Selection for Each Purpose of Use (No Suitable Format)

Analytical Instrument Data		Re-processing Only	Document-Based Compliance Review	GxP Regulations Compliance
Original Data (Instrument-Specific Format Data)		○	○	○
Numeric Data in Text Format / Data Exported in General-Purpose Formats Such as PDF		○	○	○
Derived Data	Results (e.g., peak area) processed using original data etc.	－	○	○
	Calculated results (concentration) etc.	－	○	○
	Reports including analytical conditions	○	○	○
Analytical Metadata	Sample sequence	○	○	○
	Instrument parameters	○	○	○
	Processing parameters	○	○	○
Audit Trail Metadata	Original data-related	－	○	○
	Processing results-related	－	－	○
	System-related	－	－	－

6. Data Migration Between Facilities

When a study is outsourced, the handling of study-related materials after study completion becomes an issue. In the case of paper records, transfer from the contract research organization (CRO) to the sponsor is relatively

straightforward. However, for electronic records such as measurement instrument data, various challenges arise regarding long-term archiving after study completion. Figure 7 summarizes the methods of long-term archiving of measurement instrument data at CROs.

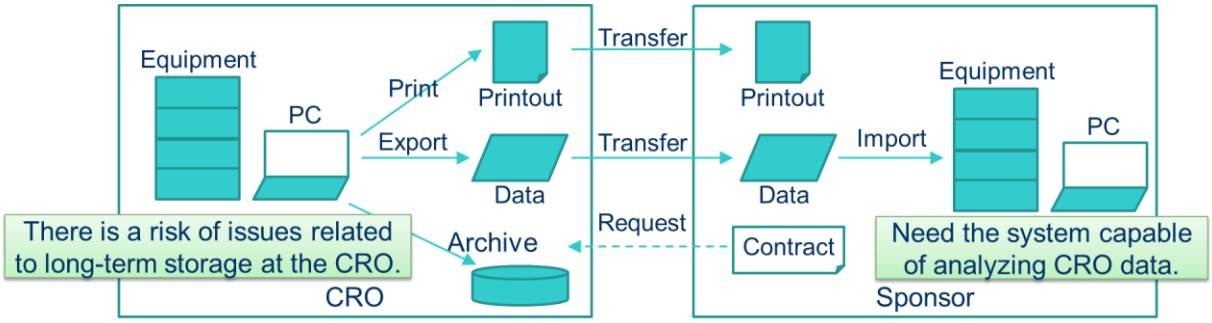

Figure 7. Long-Term Archiving of Analytical Instrument Data at CROs

One method is to print the measurement instrument data and transfer only the printed materials (upper flow in Figure 7). While this method allows for the easy return of measurement results, it has the drawback that data integrity cannot be ensured and the data cannot be used for reanalysis.

Therefore, to ensure data integrity and enable reanalysis, many facilities contract with CROs to continue preserving the measurement instrument data (see lower flow in Figure 7). With this method, data integrity is maintained and reanalysis is possible; however, the following issues may arise:

- Over time, staff members who were in place at the time of the contract may leave, resulting in unclear points of contact.
- Due to mergers or dissolution of CROs, there is a risk that materials preserved at the CRO (including measurement instrument data) will not be properly maintained.
- CROs themselves face the challenges of long-term archiving (see Sections 3.3.1 and 3.3.2).

To address the above issues, it is also necessary to consider transferring measurement instrument data from the CRO to the sponsor (see middle flow in Figure 7). By applying the methods described in Chapter 5, the reliability of this transfer process can be ensured. At a minimum, for HPLC data the measurement results returned using this method can be utilized for reanalysis and other purposes. Likewise, for other measurement instrument data, reanalysis and other uses are possible if the sponsor has an appropriate analysis application.

7. Future Challenges

To realize the method recommended in Chapter 5, the following techniques are expected to be developed.

7.1. Standardization of Data from Various Analytical Instruments

It is desirable that the long-term archiving method of analytical instrument data as shown in 5.1. above can be applied to import and use data of analytical instruments from other manufacturers. This requires manufacturers to provide compatible data formats. At present, there are many analytical instruments that don't have standard

data formats, including those shown in 3.3.4.

New data formats are developed along with new technologies. For these data formats, it is expected to provide model-independent software, output in a form compatible with existing instruments, disclose format definitions, etc.

7.2. Development of Data Package Technology

To ensure the long-term archiving method of the recommended analytical instrument data described in Section 5, the following techniques are expected to be developed. Since a packaging tool based on the specifications of the Technical Guidebook of this committee has already been made available as an open source, product development making use of this tool is also possible.

a. Technology for automatically packaging analytical instrument data
b. Technology for capturing and utilizing packaged data

8. Future Outlook

Highly improved accuracy in an analytical instrument and its increased capacity of electronic data acquired in a single analysis have resulted in an increase in the storage capacity of electronic data. In addition, a situation has occurred that the recording capacity of the electric data storage, such as the data storage server and data management system, is occupied by the data stored but hardly utilized. As a result, it is becoming a problem of cost increases related in storage of electronic data in the pharmaceutical industry.

There is a need for services in the pharmaceutical industry to contract the archiving and management of electronic data, as all companies are concerned about the storage and management of electronic data (see Section 4.2). Furthermore, a service that stores and manages the processing software of all manufacturers and provides an environment in which any manufacturer's analytical instrument data can be processed would be useful to many companies in the pharmaceutical industry.

There is also a need to provide and receive analytical instrument data and to migrate data between systems, such as through company acquisitions and alliances of relevant organizations. If services are available to retain older versions of OS and processing software on virtual environments, for many companies there are no other problems such as migrating past analytical instrument data to new systems.

The above responses seem to be technically implementable. However, for such a service to be provided as a business, the number of companies that need it and the service value paid by the company need to be large enough as a business. The investigation of the feasibility, including a survey of needs, is a future problem.

It will also be necessary to inform regulatory authorities of the status of the industry and to discuss both the technical and operational limitations. In accordance with such limitations, consideration should be given to flexibility in reviewing the required data retention periods, depending on factors such as product life cycles in the market (which generally differ significantly between pharmaceuticals and medical devices) and the importance of the preserved data (e.g., whether the data are directly related to the functionality or safety of pharmaceuticals or medical devices).

With respect to manufacturers, it is necessary for users to communicate the status of the industry and to discuss

the need for data formats. In this regard as well, support from the regulatory side is considered necessary.

Furthermore, for data mining to foster innovation, a common data infrastructure is indispensable; however, the data used therein must be reliable. If data containing errors are used for analysis, erroneous conclusions will inevitably result. To prevent this problem, it is useful, regardless of the existence of standard formats, to package the original data together with the numerical data output in a manner that ensures reliability in advance, so that the data can withstand verification at a later date.

9. Glossary

Table 6 provides a description of the terms used in this guidance.

Table 6. Terminology Description

Term	Description
Algorithm	A set of well-defined, finite rules that solves the problem by applying finite times.
Processing Parameters	Values used to set analysis conditions, such as methods for processing peak shapes, which constitute part of the analytical metadata. In the case of HPLC, these include the method of baseline setting and the method of noise reduction.
Audit Trail	A secure, computer-generated, time-stamped electronic record that allows reconstruction of the process of events associated with the creation, modification or removal of an electronic record. There are logs of original data generation, audit trails of analysis, and audit trails on the system.
QA	Quality assurance; confirmation work to certify quality.
High Performance Liquid Chromatography (HPLC):	An analytical method in which a liquid is pressurized and passed through a column, separating and detecting components (analytes) in a liquid mixture by utilizing interactions with the stationary and mobile phases.
Computerized System Validation	System that consistently fulfills the required specifications, from system design through decommissioning or migration to a new system. (ICH E6) 000230974.pdf
Sample Sequence	Injection schedule that defines the order of sample injections, injection amounts, etc., and is part of the analytical metadata.
Mass Spectrometer (MS)	A general term for instruments that perform qualitative and quantitative analysis of substances contained in a sample by ionizing molecules or atoms and analyzing their mass (m/z, which is measured) and ion quantity (the electric current generated when detecting ions).
ZIP	An archival file format that handles multiple files together as a single file. Basically, file extension "zip" is used, but there are also files which extension is not zip, such as docx.
Instrument parameters	Values used to set measurement conditions, such as the operating parameters of measuring instruments, which constitute part of the analytical metadata. In the case of HPLC, these include items such as detection wavelength and flow rate.
Time-stamp	Date and time automatically imprinted by the computer (ERES Guidelines).[24] Technique for providing information that can detect a change and verify whether there has been a change since that time.[25]
Data integrity	Completeness, consistency and accuracy of data.
Electronic signature	Signatures that are signed electronically as equivalent to handwritten signatures or seals and are constructed electronically with a set of symbols prepared, adopted, confirmed, and approved by individuals or corporations (ERES Guidelines).[24]
File configuration information in the package	A special file describing file configuration information in the package.

[24] Source: Pharmaceutical and Food Safety Bureau Notification No. 0401022 dated April 1, 2005: Use of Electromagnetic Records and Electronic Signatures in Applications for Drug Approvals or Licenses, etc. (ERES Guidelines).

[25] Time Business Forum (https://www.dekyo.or.jp/tbf/index.html) Time business lexicon.

Term	Description
Hash value	Message digest generated by the hash function.
Validation	A documented program that provides a high degree of assurance that a specific process, method, or system will consistently produce results meeting predetermined acceptance criteria. (ICH Q7) 000156438.pdf
PDF	Portable document format, a file format for the exchange of documents, including texts, images, and graphics, as specified as ISO 32000.
Bit	Smallest unit of data handled by the computer.

10. Revision History

Date	Version	Details of revision
2025.03.31	3.0	Consideration was given to methods for the long-term archiving of data from measuring instruments without appropriate standard formats (including MS), while also reflecting trends such as changes associated with the publication of technical and operational guidebooks and the use of cloud services.
2020.01.31	2.1	Revised content in preparation for English translation.
2019.04.18	2.0	Change the title from Guideline (draft) to Guidance. Reflect the solicited comments.
2018.11.28	1.0	

Appendix 1. Example of AIA Format Data

```
Netcdf demodata {
Dimensions:
        _2_byte_string = 2 ;
        _4_byte_string = 4 ;
        _8_byte_string = 8 ;
        _12_byte_string = 12 ;
        _16_byte_string = 16 ;
        _32_byte_string = 32 ;
        _64_byte_string = 64 ;
        _128_byte_string = 128 ;
        _255_byte_string = 255 ;
        Point_number = 1201 ;
        Peak_number = 4 ;
        Error_number = 1 ;
Variables:
        Float detector_maximum_value ;
        Float detector_minimum_value ;
        Float actual_run_time_length ;
        Float actual_sampling_interval ;
        Float actual_delay_time ;
        Float ordinate_values(point_number) ;
        Ordinate_values:uniform_sampling_flag = "Y" ;
        Ordinate_values:autosampler_position = "" ;
        Float peak_retention_time(peak_number) ;
        Float peak_amount(peak_number) ;
        Float peak_start_time(peak_number) ;
        Float peak_end_time(peak_number) ;
        Float peak_width(peak_number) ;
        Float peak_area(peak_number) ;
        Float peak_area_percent(peak_number) ;
        Float peak_height(peak_number) ;
        Float peak_height_percent(peak_number) ;
        Float baseline_start_time(peak_number) ;
        Float baseline_start_value(peak_number) ;
        Float baseline_stop_time(peak_number) ;
        Float baseline_stop_value(peak_number) ;
```

Float retention_index(peak_number) ;

Float migration_time(peak_number) ;

Float peak_asymmetry(peak_number) ;

Float peak_efficiency(peak_number) ;

Float mass_on_column(peak_number) ;

Char peak_name(peak_number, _128_byte_string) ;

Char peak_amount_unit(peak_number, _128_byte_string) ;

Char peak_start_detection_code(peak_number, _128_byte_string) ;

Char peak_stop_detection_code(peak_number, _128_byte_string) ;

Char manually_reintegrated_peaks(peak_number, _128_byte_string) ;

// global attributes:

:dataset_completeness = "C1+C2" ;

:aia_template_revision = "1.0.1" ;

:netcdf_revision = "VERSION of Aug 26 2015 17:18:35 $" ;

:languages = "English" ;

:administrative_comments = "" ;

:dataset_origin = "Supplier" ;

:dataset_owner = "" ;

:dataset_date_time_stamp = "20030116164231+0900" ;

:injection_date_time_stamp = "20030116164231+0900" ;

:experiment_title = "" ;

:operator_name = "Operator" ;

:separation_experiment_type = "" ;

:company_method_name = "" ;

:company_method_id = "" ;

:pre_experiment_program_name = "" ;

:post_experiment_program_name = "" ;

:source_file_reference =

"C:¥¥Data¥¥Project1¥¥Demo_Data-001.dat" ;

:sample_id_comments = "This is comment."/;//Data File Comment

Sample_id = "1";//Sample ID

::sample_name = "STD";//Sample name

::sample_type ="Standard";//sample type

:sample_injection_volume = 10.f ;

:sample_amount = 1.f ;

:detection_method_table_name = "" ;

:detection_method_comments = "This is comment.";// Comment in Data File.

:detection_method_name =

"C:¥¥Sample¥¥LC¥¥Demo_Method.met" ;

: detector_name = "detector A";

Units of::detector_unit ="Volts";//detector

:raw_data_table_name =

"C:¥¥Data¥¥Project1¥¥Demo_Data-001.dat" ;

::retention_unit ="Seconds";//time units

:peak_processing_results_table_name = "" ;

:peak_processing_results_comments = "This is comment.";// Comment in Data File.

:peak_processing_method_name =

"C:¥¥Sample¥¥LC¥¥Demo_Method.met" ;

:peak_processing_date_time_stamp = "" ;

Data:

Detector_maximum_value = 0.01606795 ;

Detector_minimum_value = -0.002146373 ;

Actual_run_time_length = 600 ;

Actual_sampling_interval = 0.5 ;

Actual_delay_time = 0 ;

Ordinate_values = -3.786087e-006, -3.786087e-006, 6.341934e-007 ,

2.627373e-006, 4.763603e-006, 6.165505e-006, 7.691383e-006 ,

.....

5.438328e-005, 5.547047e-005, 5.688667e-005, 5.883217e-005 ,

6.048202e-005, 6.172657e-005, 6.249905e-005, 6.396294e-005 ,

6.585121e-005, 6.725788e-005, 6.734371e-005, 6.756782e-005 ,

6.769657e-005, 6.738186e-005, 6.752014e-005, 6.912708e-005 ,

7.059574e-005 ;

Peak_retention_time = 159.271, 194.451, 237.258, 278.901 ;

Peak_amount = 1, 1, 1, 1 ;

Peak_start_time = 151.5, 180.5, 229.5, 271.5 ;

Peak_end_time = 169.5, 204, 246, 287.5 ;

Peak_width = 18, 23.5, 16.5, 16 ;

Peak_area = 73509.84, 83363.7, 53565.35, 57335.54 ;

Peak_area_percent = 27.45215, 31.13206, 20.00391, 21.41188 ;

Peak_height = 13039.91, 15250.72, 11457.2, 13248.39 ;

Peak_height_percent = 24.60536, 28.77699, 21.6189, 24.99875 ;

Baseline_start_time = 151.5, 180.5, 229.5, 271.5 ;

Baseline_start_value = 4.217148e-005, 0.0005049753, 0.00154314 ,0.002390246 ;

Baseline_stop_time = 169.5, 204, 246, 287.5 ;

Baseline_stop_value = 0.0003286457, 0.001030946, 0.00189703, 0.0025459 ;

Retention_index = 0, 0, 0, 0 ;

Migration_time = 0, 0, 0, 0 ;

Peak_asymmetry = 1.193704, 1.139562, 1.168244, 1.165645 ;

Peak_efficiency = 4246.021, 6958.823, 12828.54, 20710.6 ;

Mass_on_column = 0, 0, 0, 0 ;

Peak_name =

"Methyl paraben", "Ethyl paraben", "Propyl paraben", "Butyl paraben" ;

Peak_amount_unit = "", "", "", "" ;

Peak_start_detection_code = " ", " ", " ", " " ;

Peak_stop_detection_code = " ", " ", " ", " " ;

Manually_reintegrated_peaks = "", "", "", "" ;

}

Appendix 2. Explanation of the Purpose of Using Analytical Instrument Data

This guidance provides examples of the selection of measurement instrument data for different purposes of use in Tables 4 and 5; however, in actual operations, each company is expected to make its own judgment according to the circumstances.

The rationale for the categorization in Table 4 is provided here for reference.

Table 4. (Reshown) Example of Selection of Analytical Instrument Data for Each Purpose of Use

Analytical Instrument Data		Re-processing Only	Document-Based Compliance Review	GxP Regulations Compliance
Original Data (Instrument-Specific Format Data)		○	○	○
Standard Format (AIA) Data		○	○	○
Derived Data	Results (e.g., peak area) processed using original data etc.	－	○	○
	Calculated results (concentration) etc.	－	○	○
	Reports including analytical conditions	○	○	○
Analytical Metadata	Sample sequence	○	○	○
	Instrument parameters	○	○	○
	Processing parameters	○	○	○
Audit Trail Metadata	Original data-related	－	○	○
	Processing results-related	－	－	○
	System-related	－	－	－

○: Storage －: No Storage Required

Re-processing Only

One of the objectives of the company to store analytical instrument data for long periods of time is to enable the possibility of re-processing as needed at a later date. Requirements for re-processing were set as follows and are reflected in Table 4.

- It is imperative that electronic data can be imported into the software used for re-processing. Since there is a possibility that the original data cannot be imported, its standard format data is also stored with it.
- Analytical metadata is stored as a source of sample identification and analytical conditions.
- The package is timestamped to prove that it has not been changed before re-processing.

Document-Based Compliance Review

At the time of application for new drug approval, it is investigated whether pharmacological, pharmacokinetic,

clinical pharmacokinetic, and quality studies have been collected and prepared in accordance with the “Standards of Reliability of Application Data (Articles 43, 61, 114-22, 114-42, 137-25, or 137-42 of the Enforcement Regulations of the Pharmaceutical and Medical Devices Act)” in a document-based compliance review. It is natural to store data from studies conducted in accordance with the above criteria, but there are differences in the long-term archiving methods and concepts among institutions. Table 2 requirements were established based on the following considerations.

- Once inspection by the internal audit department such as QA/QC is completed, the original data, derived data, analytical metadata, and audit trail metadata related to the original data are stored in the package as it is.
- In addition, standard format data is also archived in the package so that dynamic data such as waveforms can be reproduced even after long-term archiving.
- The package will be timestamped to prove no change after in-house audit.

GxP Regulations Compliance

Although it is imperative to store study materials in accordance with GxP regulations, there are differences in long-term archiving methods and perceptions among institutions. Table 2 requirements were established based on the following considerations.

- Once inspection by the internal audit department such as QA/QC is completed, the original data, derived data, analytical metadata, and audit trail metadata related to the original data are stored in the package as it is.
- In addition, standard format data is also archived in the package so that dynamic data such as waveforms can be reproduced even after long-term archiving.
- The package will be timestamped to prove no change after in-house audit.

In addition, audit trails of systems such as log-in and log-out were not included in the package for the following reasons:

- Although the package proposed in this guidance is made in an analytical run, the system's audit trails are managed throughout the system's use rather than in an analytical run.
- It is because the system's audit trail is used to detect inappropriate activities other than packaged data. (In GxP facilities, inappropriate operations other than data are confirmed through regular audits by internal auditing departments such as QA/QC.)

R&D Data Archiving Committee Guidance Preparation Members

Committee Member	
Yasushi Okada	Ricoh Japan Co., Ltd.
Sayuri Uehara	Astellas Pharma, Inc.
Akira Yamazaki	Kyowa Kirin Co., Ltd.
Satoshi Morino	Eisai Co., Ltd.
Naoto Miyaji	Lang Edge, Inc.
Kozo Ogimoto	
Yoshiaki Hiraishi	Terumo Corporation
Junpei Miura	Business Engineering Corp.
Yoshimasa Haruyama	Business Engineering Corp.
Yoshiaki Yamada	Hitachi High-Tech Science Corporation
Michihiro Kimura	Japan Image and Information Management Association
Collaborating Members Japan Society of Quality Assurance (JSQA)	
Hiroyasu Ishiguro	Lead Chemical Co., Ltd.
Akiyoshi Otake	Astellas Pharma, Inc.
Yoshikazu Masaki	Otsuka Pharmaceutical Factory, Inc.