

測定機器データの長期保存ガイダンス

第 2.1 版

2020 年 1 月 31 日



公益社団法人日本文書情報マネジメント協会
R&D データ保存委員会

目次

1. 目的	3
2. 適用範囲	3
3. イントロダクション	3
3.1. 測定機器データの概要	3
3.2. 紙に印刷した測定機器データの問題点	4
3.3. 測定機器データの懸念点	5
3.3.1. 再解析	5
3.3.2. 長期保存	5
3.3.3. データインテグリティ	6
3.3.4. 測定機器データのフォーマット	8
4. 方針	9
4.1. 解析結果に関する考え方	10
4.2. 長期保存に関する考え方	10
4.3. データインテグリティに関する考え方	11
5. 推奨する測定機器データ及び解析レポートの長期保存方法	11
5.1. 測定機器データの長期保存方法	12
5.2. 解析レポートの長期保存方法	15
6. 今後の課題	15
6.1. 各種測定機器データの標準化	15
6.2. データパッケージ技術の開発	15
7. 将来の展望	16
8. 用語集	17
9. 改訂履歴	18
付録1 AIA形式データの例	19
付録2 表2の詳細説明	23

1. 目的

本ガイダンスの目的は、製薬業界の研究施設、試験検査施設等において、各種の測定機器から出力された電子データ（以下、「測定機器データ」という）を、そのデータが再解析される可能性があることを前提として信頼性を保ったまま長期間に渡り安心して保存管理する方法を提示することにある。

2. 適用範囲

本ガイダンスは、製薬業界において以下の目的で取り扱われる測定機器データを対象とする。

- a. 規制当局への申請及び報告に用いるデータ
- b. GxP 規制¹で保存が義務付けられているデータ
- c. 施設において保存の必要があると考えられているデータ

なお、測定機器データは、再解析される可能性があることを前提としている。もし天秤や pH 計などの測定機器から出力されたデータをそのままの状態で確定し、再解析を行わない場合は、本ガイダンスの対象には含まれない。

3. イントロダクション

IT 環境の急激な変化に伴い、保存しなければならない電子記録の量は急速に増加している。その一方で、紙に印刷した記録と異なり、電子記録は新しいシステムへの移行やソフトウェアのバージョンアップなどによって、将来的に読めなくなることが危惧されている。

また、試験を外部委託した場合、委受託試験における施設間の測定機器データの引き渡しに際して、紙に印刷した測定機器データであれば印刷物を施設間で容易に引き渡すことができる。しかし、測定機器データそのものの場合、真正性²を保った状態で、どのような方法で引き渡すのか、引き渡された測定機器データをどのように保存するのかなどについて明確な基準がなく、施設が各々最適と考える方法で対応せざるを得ない状況である。

製薬業界においては、測定機器データを信頼性が保たれ、再解析ができる状態で長期間保存しなければならないため、上記の課題はより重大である（3.3 参照）。

本ガイダンスの策定に至った前提や背景を以下に示す。

3.1. 測定機器データの概要

測定機器、又は測定機器に接続されたコンピュータには、測定データや解析結果のほか、測定条件、解析条件、サンプルスケジュール、電子署名及び監査証跡、並びにシステムのログなどが保存

¹ 患者の安全性や試験の信頼性を確保することを目的に規制当局で定められた基準の総称。製薬業界の代表的な法令を以下に例示するが、これに限らない。

GLP (Good Laboratory Practice) : 「医薬品の安全性に関する非臨床試験の実施の基準」

GCP (Good Clinical Practice) : 「医薬品の臨床試験の実施の基準」

GMP (Good Manufacturing Practice) : 「医薬品の製造管理及び品質管理の基準」

GVP (Good Vigilance Practice) : 「医薬品の製造販売後安全管理の基準」

² 参照：薬食発 第 0401022 号平成 17 年 4 月 1 日 医薬品等の承認又は許可等に係る申請等における電磁的記録及び電子署名の利用について（ERES 指針）、3.1.1.電磁的記録の真正性の項

される。最近ではネットワークを介してデータ保存サーバ、データ管理システム等に保存されることも多い。

本ガイダンスでは、これら測定機器データについて以下の用語を用いる。

測定機器データの保存形式としては、下記 A., B.の2つがあり、測定機器データには、下記 a., b., c.の3種類のデータが含まれる。

A. オリジナルデータ

測定機器から出力された電子データ。オリジナルレコードともよばれ、GxP 規制下では生データの一部とされる。機器固有のフォーマットで出力されることが多い。機器によって、測定データ以外に派生データとメタデータを含む場合と含まない場合がある。

B. 標準フォーマットデータ

AIA などの標準フォーマットに変換した測定データ及び一部の派生データとメタデータ。残りの派生データとメタデータが変換されずに欠落するが、欠落した分を補完する資料等と合わせて保存することによって、変換前のデータと同等と見なすことができる。

a. 測定データ

測定機器で得られたデータ。高速液体クロマトグラフ (HPLC) などのクロマトグラムデータ、赤外分光光度計 (IR), 核磁気共鳴装置 (NMR), 質量分析計 (MS) などのスペクトルデータ、天秤の秤量データなど。

b. 派生データ

オリジナルデータを使って解析された結果 (ピーク面積等) や計算結果 (濃度等) など。

c. メタデータ

データの意味を明確にするために補足される単位等の情報、データの取得された日時や測定機器等を特定する情報、測定条件、監査証跡など。FDA はデータインテグリティのガイダンス³において、メタデータを具体的に説明している。本ガイダンスでは FDA の考え方を参考にしつつ、測定機器データに関連するデータと、監査証跡を分けて示す。即ち、測定条件、解析条件、サンプルスケジュールなどを分析メタデータと表記し、監査証跡を監査証跡メタデータと表記する。

3.2. 紙に印刷した測定機器データの問題点

従来、日本の製薬業界では、測定機器データや解析レポートを印刷して紙で保存することが一般的に行われてきた。しかし、次の問題点を解決し、データのインテグリティを確保するために、最近では電子データを保存することが求められている。

a. 印刷されたデータは再解析できない。

b. 電子データの中に含まれていても通常印刷されないデータ (解析条件や測定条件等)、印刷できないデータ (3次元データ等) は全て失われてしまう。

c. 紙を保存する膨大なスペースが必要になる。

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

3.3. 測定機器データの懸念点

3.3.1. 再解析

製薬業界では測定機器データを取得してから数年後に再解析しなければならないことがある。例えば類縁物質試験において、あるロットで問題となった類縁物質が、過去のロットにどの程度含まれていたかを調査する場合がある。その際、未報告の微小ピークを調査するために、過去の測定機器データの再解析が必要となることがある。

この時、適切に保存された測定機器データがあり、再解析できる環境が用意できたとしても、同じメーカーのソフトウェアで再解析した場合を除き、アルゴリズムの違いから、同一の結果が出ることは保証できない。

また、同一メーカーのソフトウェアであっても、バージョンアップされている場合、元の解析結果と同一の結果が得られるとは限らない。

この理由には、次のものがある。

- a. 解析アルゴリズムを変更している。(計算精度の向上も含む)
- b. 計算機内部では 10 進数ではなく 2 進数の有限桁数で計算しているため、最小桁には誤差が含まれており、計算結果にも誤差を含む。
- c. コンピュータも現在は 32 ビットから 64 ビットに変化しつつあるので、これにより差が生じる可能性がある。

これらを避けるには、再解析が可能な環境として、解析ソフトウェアだけではなく OS まで含めたコンピューター式を保管することが必要となるが、現実的ではない。

3.3.2. 長期保存

再解析の必要性を考慮した場合、製薬業界では測定機器データを 30 年位再解析できる状態で保存することが望まれる⁴。しかし、ハードウェアの更新やソフトウェアのバージョンアップが行われた場合、解析ソフトウェアが新しい OS に対応できず、長期間にわたって保存された測定機器データが利用できない恐れがある。また、利用しているデータベースソフトウェアがサポートを終了したとき、他メーカーのシステムに測定機器データを移行できない恐れもある。

以下、代表的な対応策とそれぞれの懸念点について記す。

A) 解析用コンピュータを物理的に保存

懸念点：

- a. 管理対象システムの増加
- b. 機械的寿命 (コンピュータ本体, プリンター, 外部ストレージ)
- c. 維持費 (ライセンス費含む)
- d. 保存場所
- e. 定期点検
- f. 電子データのバックアップの維持

⁴ 例：GMP 省令第 20 条第 3 号、第 30 条など。

- g. 旧システムを操作できる作業者の確保

B) 解析用コンピュータを仮想化し保存

懸念点：

- a. 管理対象システムの増加
- b. 維持費（ライセンス費含む）
- c. 定期点検
- d. 電子データのバックアップの維持
- e. 旧システムを操作できる作業者の確保
- f. ハードウェアキー⁵、インタフェースカード⁶等のハードウェアが必要な場合の追加措置
- g. 仮想化の対象となるソフトウェアのライセンス形態⁷
- h. 完全に互換性がとれているかの検証が必要

C) 現行システムから後継システム（又は互換システム）への移行

懸念点：

- a. 移行可能なシステムの制約
システム導入時にその制約を知ることは困難である。
- b. 移行可能な範囲の制約
例えば、監査証跡は移行不可である。

D) 標準フォーマットを介したデータ移行

懸念点：

- a. 自動・手動を問わず、移行元のシステムにデータとメタデータをエクスポートできる機能が、移行先のシステムにそれらをインポートできる機能が必須である。
- b. メタデータとのリンクの喪失、又はリンクの移行が困難である。
- c. 移行元のシステムと移行先のシステムとのデータ管理水準の違いにより、データインテグリティの保持が困難である。例えば、そのシステムで生成したデータ以外には監査証跡が記録されない。

本ガイダンスでは、「D) 標準フォーマットを介したデータ移行」を推奨する。推奨する理由は4章で述べる。

3.3.3. データインテグリティ

再解析を行うことを前提に、測定機器データを保存したり、保存場所を移動させたりする際には、その信頼性が保たれていることが重要である。特に、規制当局への申請及び報告に用いる資料

⁵ 不正アクセスや不正コピーを防止する目的で、ソフトウェアなどの利用制御に用いられる USB デバイスなどのハードウェア（dongle）。

⁶ コンピュータに入出力インターフェースを追加する拡張カード。

⁷ ユーザーではなくハードウェアにライセンスが必要な場合がある（Microsoft Windows DSP 版、OEM 版など）。

に用いる測定機器データについては、データインテグリティの確保を求められるようになってきている。このデータインテグリティ確保のためには、測定データだけでなく派生データ、メタデータ（分析、監査証跡）も含めて、判読性を保った状態でセキュリティ管理のもと保存することが要求されているが、要求を満たすには、3.3.1, 3.3.2の懸念点を含め、困難な点がある。

下記に規制当局の要求事項が書かれたガイダンス等の抜粋を示す。

電子データの保存に対する規制当局の要求

FDAは、“Questions and Answers on Current Good Manufacturing Practices, Good Guidance Practices, - Records and Reports”というガイダンス⁸の中で、次のようなQ&Aを掲載している。なお、回答については、関連する部分のみ抜粋し、理解しやすいように整理した。

Q) 医薬品の製造と試験で使用されるラボのコンピュータ化システムが生成する電子記録及び関連する印刷されたクロマトグラムに対して、Part 11やプレディケートルール要件 (21 CFR Part 11) はどの様に適用されるのか？

- A) 印刷されたクロマトグラムは、§ 211.68で要求されているように、クロマトグラムを作成するために使用された電子データの正確かつ完全なコピーとは見なされない。クロマトグラムには、一般に、例えば注入シーケンス、装置メソッド、解析メソッド、監査証跡などは含まれていない。これらのすべてがクロマトグラムの作成に使用され、信頼性と関連付けられている。したがって、
- a. 医薬品の製造と試験で使用される印刷されたクロマトグラムは、Part 211におけるプレディケートルール要件を満たしていない。
 - b. ラボのコンピュータ化システムによって生成された電子記録は、これらの要件に則って維持されなければならない。

FDAは、また、次のようなWarning Letter⁹を出している。

- a. 会社は、HPLCテストをサポートするすべての電子生データを削除しました。さらに、会社は、注入シーケンス、機器メソッド、テストの統合メソッドなどの基本的なクロマトグラフィ情報を保持していませんでした。
- b. システムでは、オペレーターがファイルを削除できました。この行為を制御するためのアクセス管理手順がありませんでした。

FDAのガイダンス³は、データインテグリティのためにALCOAを要求しており、PIC/Sのガイダンス¹⁰は、それに加えてCCEAを要求している。日米欧の規制当局の期待に応じ、当局による査

⁸ 出典：<https://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm124787.htm#3>

⁹ 出典：<https://www.fda.gov/ICECI/EnforcementActions/WarningLetters/2014/ucm404316.htm>
文書番号 320-14-10
<https://www.fda.gov/ICECI/EnforcementActions/WarningLetters/2017/ucm546483.htm>
文書番号 320-17-29

¹⁰ 出典：PIC/S Guidance Good Practices for Data Management and Integrity in Regulated GMP/GDP Environments (PIC/S, PI 041-1 (Draft 3) 30 Nov. 2018)

察（適合性書面調査など）に耐えうるものとするためには、これらをまとめた ALCOA+の要求を満たす方法を確立する必要がある。

ALCOA

Attributable（帰属性）：署名や捺印等により帰属や責任の所在が明らかである
Legible（判読性）：誰もが読める字で簡潔かつ明瞭に記録されている
Contemporaneous（同時性）：発生と同時にあるいは遅滞なく記録されている
Original（原本性）：記録内容が原本として管理されている
Accurate（正確性）：事実が正確かつ客観的に記録されている

CCEA

Complete（完全性）：記録に欠損がなく完結している
Consistent（一貫性）：記録（データ）が理にかなない矛盾がない
Enduring（耐久性）：記録の保存が永続的である
Available（利用可能性）：記録を必要な時に取り出せる

また、MHRA のガイダンス¹¹は、6.11.2 で真正なコピー（True copy）について以下のように定めている。

- a. 必要ならば、真正なコピーはオリジナルレコードとは異なる電子ファイル形式で保管してもよいが、データの全ての意味が保たれ、履歴が再構築できることを保証するためにメタデータと監査証跡を保持しなければならない。
- b. オリジナルレコードとその真正なコピーは、記録の完全性を保持しなければならない。当該コピーの完全性を検証して記録する文書化されたシステムが確立されているなら、オリジナルレコードの真正なコピー（例えば、紙記録のスキャンしたもの）を、オリジナルレコードの代わりに保存してもよい。組織は、オリジナルレコードの破棄に伴うあらゆるリスクについて考慮すること。
- c. レビュー、バックアップ、アーカイブなどの目的のために、関連するメタデータを含めた電子データの真正なコピーを作成できるようにすること。真正なコピーであることを保証するために、正確で完全なコピーは、データの意味（例えば、日付フォーマット、前後関係、レイアウト、電子署名及び承認）と完全な監査証跡を含むこと。保存期間を通じて、真正なコピーの動的な機能が維持されるよう考慮すること。

3.3.4. 測定機器データのフォーマット

測定機器は様々な種類のデータを発生させる。あるメーカーの測定機器のオリジナルフォーマットで保存した電子データを、そのフォーマットのまま他のメーカーの測定機器に移行することは難しい。

そのため、メーカーに依らない標準フォーマットの検討もなされており、ここでは、よく使用さ

¹¹ 出典：‘GXP’ Data Integrity Guidance and Definitions (MHRA, Revision1 March 2018)

れるクロマトグラフデータを例として取り上げる。クロマトグラフデータの標準フォーマットとしては、次に挙げるフォーマットが提唱されているが、それぞれに一長一短があり、統一できていないのが現状である。

a. AIA (Analytical Instrument Association)

米国の分析機器工業会 (Analytical Instrument Association - AIA) が定めた、HPLC 等のクロマトグラフデータの標準フォーマット。バージョンが違くと読み出せない事がある。

b. NetCDF (Network Common Data Form)

大気研究大学連合 (University Corporation for Atmospheric Research - UCAR) Community Programs (UCP) の一部である Unidata が開発した、科学的なデータを表現する為の機種に依存しないデータフォーマット。

c. JCAMP (Joint Committee on Atomic and Molecular Physical Data)

国際純正・応用化学連合 (International Union of Pure and Applied Chemistry - IUPAC) が引き継いだ NMR, IR, MS 等のスペクトルデータの標準フォーマット。

d. AnIML (Analytical Information Markup Language)

アメリカ材料試験協会 (American Society for Testing Materials) の XML 標準フォーマット。まだ最終化されていない XML 形式のフォーマットで、データ量が大きくなることと (オリジナルデータの 10~20 倍)、機器のデータの取り込み速度が速くなっているのにもかかわらず書き込み速度が追いつかない欠点があるため、標準化が進んでいない¹²。

なお、上記いずれのフォーマットも分析メタデータの全てを含んでいる訳ではない。

クロマトグラムデータやスペクトルデータは標準フォーマットに変換が可能であるが、測定条件、解析条件などの分析メタデータは測定機器メーカー各社で仕様が異なる。また、解析アルゴリズムは各メーカーで異なるため、解析パラメーターが移行されたとしても、他メーカーのデータを完全に再現することはできない。しかしながら、定量分析の場合は、標準品に対する相対値がほぼ同じであれば、他社の機器でも同等な定量値を求めることができる。このような定量分析の例として、標準品を用いて作成された検量線による濃度算出が挙げられる。

また、標準フォーマットといっても、メーカーごとに取り込むメタデータは異なっており、フォーマットとして規定されている項目にデータが入っていないケースや、単位や時刻などに関するデータの形式が違うこともある。

4. 方針

本章では、5章で述べる、推奨する測定機器データの長期保存方法の背景となる考え方を示す。表1に、3.3で挙げた懸念点に対する対応方針のまとめを示す。

¹² XML形式のフォーマットは、通信のアルゴリズムを改良することで、通信速度が向上する可能性があるが、基本は機械語で組み込んでいるため、通信アルゴリズムの修正は困難。

表1 懸念点に対する対応方針（まとめ）

	懸念点	対応方針
1	再解析時に、解析アルゴリズムは各メーカーで異なるため、解析パラメーターが移行されたとしても、他メーカーのデータを完全に再現することはできない。(3.3.1)	4.1 に示す
2	再解析時に、同一メーカーであっても、ソフトウェアのバージョンの違いにより、元の解析結果と同一の結果が得られるとは限らない。(3.3.1)	4.1 に示す
3	長期保存のためにデータを移行する場合、オリジナルデータとメタデータとのリンクが喪失する。又は、リンクの移行が困難である。(3.3.2)	4.2 に示す
4	長期保存のためにデータを移行する場合、移行先のシステムでのデータ管理水準には制限があり、データインテグリティの維持が困難である。例えば、移行先のシステムで生成したデータ以外には監査証跡が記録されない。(3.3.2)	4.3 に示す

4.1. 解析結果に関する考え方

3.3.1 で述べたように、測定機器データがあっても、元の解析結果と同一の結果が得られるとは限らない。オリジナルデータを完全に保存できても、解析する時に入手可能なコンピュータ、ソフトウェアにより、微妙な差が含まれることは許容せざるを得ない。しかし、差は小さいので、実質上、試験の目的に大きな影響を与えない場合が多いと考えられる。

従って、データ長期保存時に何をどこまで求めるのかを明確にすることにより、必要以上に要求水準を高めないことが必要である。想定される要求水準としては以下のようなものが考えられる。

- a. 再解析結果が元の解析結果と一致している必要がある。
- b. 再解析することができれば、アルゴリズム変更やその他の理由による解析結果の違いは許容される。
- c. 前回と別の観点から見直すことができれば良い。例えば類縁物質の存在が懸念される部分を拡大してピークを確認する必要があるなど。

4.2. 長期保存に関する考え方

本ガイダンスは、測定機器データを再解析可能な状態で保存することが期待される期間として、10年から30年を想定している。このため、本ガイダンスでは、以下の方法を提案する（図1）。

- a. 測定機器（現行システム）から、オリジナルデータや派生データのほか、標準フォーマットデータ、分析メタデータ（測定条件、解析条件、サンプルスケジュール）、監査証跡メタデータといった測定機器データをエクスポートしてパッケージ化する。
- b. このパッケージをストレージ（データ保存サーバ、データ管理システム等）に保存し維持する。
- c. 再解析等が必要になった際に、このパッケージを新たな解析環境（後継システム）にインポートして使用する。

これにより、測定機器データの測定機器からの独立が図られ、流動化が促進されると共に、解析環境を提供する受託サービスが実現される可能性が生まれる（7.参照）。

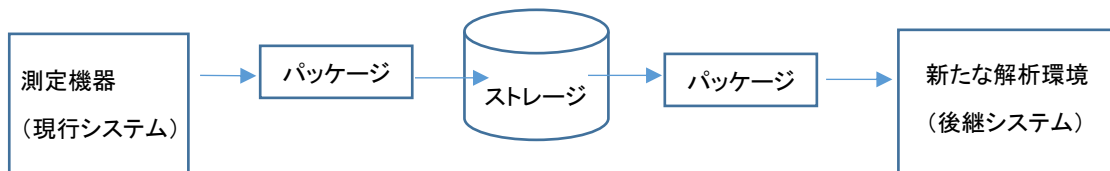


図 1 データのパッケージ化

4.3. データインテグリティに関する考え方

データインテグリティは、オリジナルデータ、派生データ、標準フォーマットデータ、分析メタデータ、監査証跡メタデータ、データプロセス、OS/ハードウェアと密接に関係している（図 2）。

エクスポートしてパッケージ化した測定機器データには、エクスポート・パッケージ化作業を行う時点までの測定機器データのみが含まれる。

本ガイダンスでは、監査者が所定の条件下でデータインテグリティが確保されていることを確認することを条件に、測定機器データをエクスポートすることを推奨する。エクスポート・パッケージ化作業は、測定・解析終了直後、ソフトウェアバージョンアップ時など、複数のタイミングで実行される。

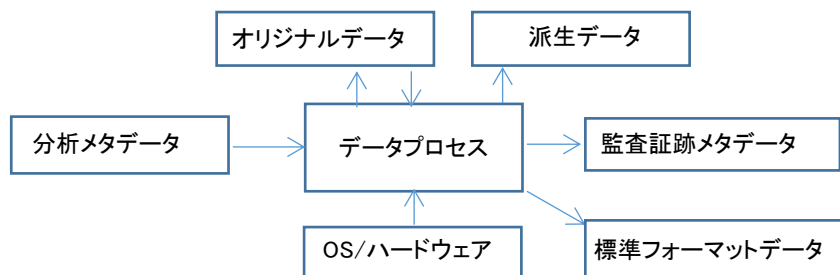


図 2 データインテグリティに係る要素

測定機器データをエクスポートした後は条件付きでデータインテグリティが確保されることになるが、データをエクスポートしてから新たなシステムにインポートするまでのプロセスの適切性を第三者（QA など）が保証すれば、たとえデータのエクスポートやインポートが自動化されていなかったとしてもデータの信頼性を確保できるものと考えられる。

5. 推奨する測定機器データ及び解析レポートの長期保存方法

再現性の確保に必要な測定機器データをパッケージとして一つにまとめ、このパッケージに対して改変を防止する仕組みを整えることが本ガイダンスにおける測定機器データの長期保存に関する

基本的な考え方である。

なお、測定に際しては、測定機器データ以外に解析結果をまとめた解析レポートも作成され、測定機器データと共に解析レポートも長期間保存が必要になる。測定機器データはそのままの状態でも長期間保存しなければならないが、解析レポートは後日修正されることがある。そのため、測定機器データの長期保存方法と解析レポートの長期保存方法については分けて記載することとする。

測定機器データには様々な種類がある。本ガイダンスでは、測定機器データの代表例としてクロマトグラフデータを、その標準フォーマットの代表例として AIA フォーマット¹³を取り上げて、測定機器データや解析レポートの長期保存の方法を検討した。

5.1. 測定機器データの長期保存方法

高速液体クロマトグラフ (HPLC) 等から出力されるクロマトグラフデータの AIA フォーマットには、クロマトグラムデータ以外の派生データやメタデータがすべて含まれているわけではない。付録に、AIA フォーマットデータの例を示す。一方、サンプルスケジュールは重要な情報であり、PDF ファイルとして出力できるが、標準フォーマット (AIA) データファイルと関連付けられない。これらのデータの間に関連付けについては、複数のファイルをひとかたまりにして保存することが有効である。

なお、HPLC 測定の場合には、他の目的のために準備された試料を主目的のために準備された試料と一緒に分析する場合もある。このような状況で一連のデータを保存するときに必要なデータのみ抽出すれば、恣意的な抽出であると疑われる。そのため、他の目的のために測定したデータも含めて全て保存すべきである。

測定機器データパッケージのイメージを図 3 に示す。測定機器データパッケージには、オリジナルデータと、測定データの動的な状態を維持して長期間保存できる標準フォーマット (AIA) データが含まれる。そのパッケージに派生データ、分析メタデータ、監査証跡メタデータなどを加えることによって、測定機器データの真正性を確保する。

これら複数のファイルの入れ物 (パッケージ) として、ZIP パッケージを利用することが推奨される¹⁴。

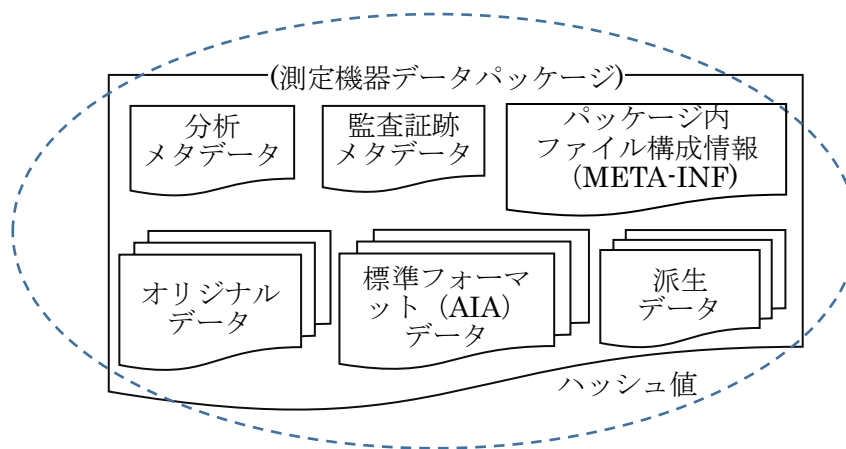
ZIP パッケージには、オリジナルデータ、標準フォーマット (AIA) データ、派生データ、メタデータに加えて、パッケージ内のファイル構成情報を記述した特別なディレクトリ (META-INF)

¹³ 日本ウォーターズ株式会社、株式会社島津製作所、株式会社日立ハイテクサイエンスの HPLC において、他社が作成した AIA フォーマットのデータを読み込むことで波形を再現し、自社の解析ソフトにて取り扱えることが確認されている。(株式会社日立ハイテクサイエンスの HPLC は AIA フォーマットの読み込み機能がないため、書き出しのみ確認された。)

¹⁴ そのほかのパッケージ候補として、ZIP ベースの目的別パッケージ、PDF、XML 構造体が挙げられる。ZIP ベースの目的別パッケージは、既に様々な形式が存在する (例えば、office 文書の OPC、電子図書の EPUB、欧州で標準化が進められている ASiC (Associated Signature Container, ETSI TS 102 918) など)。これらは、パッケージ内の META-INF の定義が違うだけとみなすこともできる。将来的には、測定機器データ専用の META-INF を定義することも候補となる (この場合はそれに対応する新たな拡張子を登録することになる)。

を格納する。この META-INF は、ディレクトリ名により識別することになる。なお、META-INF は、パッケージの外側で管理することも可能である。この場合は、パッケージのハッシュ値でリンク付けを行う。ハッシュ値の生成には、例えばタイムスタンプあるいは電子署名といった適切な技術を用いる。タイムスタンプを用いる場合は、タイムスタンプ局の証明書の有効期間満了後の有効性保証の方法を確立する必要がある。その方法の一つとして、新聞や雑誌等に公告し、見読性確保のために国立国会図書館を利用することが考えられる。

再解析を行う場合は、パッケージ化されたデータを直接操作するのではなく、データをコピーし、このコピーを使って再解析を行うことで、オリジナルデータの誤消去や、誤って上書きしてしまう危険性を排除する。



破線がハッシュ演算の範囲。ハッシュ値はパッケージとは別に管理される。

図 3 測定機器データのパッケージ化

データ保存に対する要件は、利用目的や利用状況、及び要求される信頼性のレベルや保存期間によって異なる。利用目的には、再解析のみできる、適合性書面調査で調査の対象となる、GxP 規制への準拠が求められるなどが含まれる。表 2 にパッケージに格納すべき測定機器データの一例を示す。(詳細説明は付録 2 を参照)

表2 利用目的ごとの測定機器データの選択の例

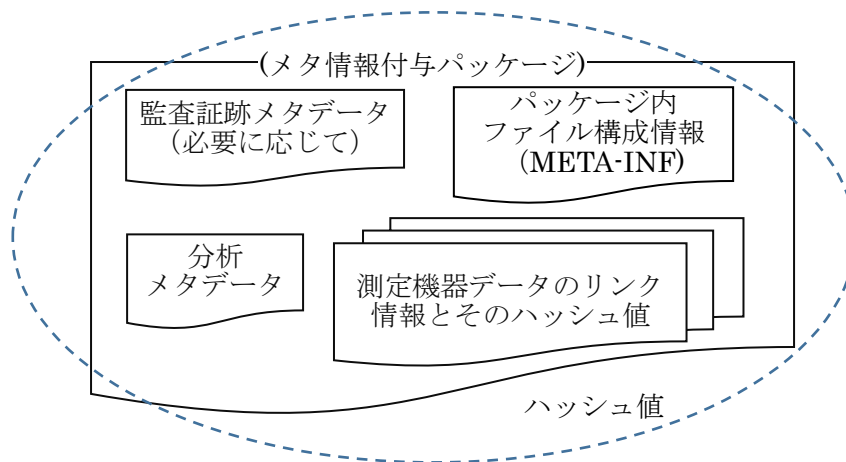
測定機器データ		再解析のみ	適合性 書面調査	GxP 規制 準拠
オリジナルデータ(機器固有フォーマットのデータ)		○	○	○
標準フォーマット (AIA) データ		○	○	○
派生データ	オリジナルデータを使って 解析された結果 (ピーク面積等) など	—	○	○
	計算結果 (濃度等) など	—	○	○
分析メタデータ	サンプルスケジュール	○	○	○
	装置パラメーター	○	○	○
	解析パラメーター	○	○	○
監査証跡メタデータ	オリジナルデータに関連するもの	—	○	○
	解析結果関連	—	—	○
	システム関連	—	—	—

○：格納する —：格納不要

分析メタデータを個々の測定機器データパッケージに含める代わりに、図4のようなメタ情報付与パッケージを構成して、分析メタデータを外部から付与することも可能である。

なお、メタ情報付与パッケージの中の「測定機器データのリンク情報とそのハッシュ値」には、分析メタデータを含まない測定機器データパッケージの存在場所を示したリンク情報とそのハッシュ値を格納する。

これにより、他の目的ですでにパッケージ化されている電子データを参照することと保存する電子データの容量を削減することが可能になる。



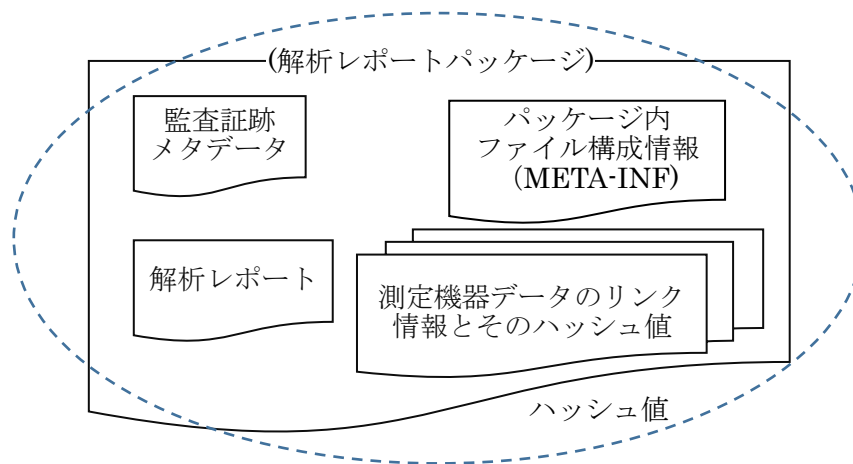
破線がハッシュ演算の範囲。ハッシュ値はパッケージとは別に管理される。

図4 分析メタデータのパッケージ化 (メタ情報付与パッケージ)

5.2. 解析レポートの長期保存方法

解析レポートは、測定機器データの解釈を変更したり、後日異なる角度から再解析を行ったり、別の解析結果との比較を行ったりするなどによって、修正することもある。その真正性を確保するために、試験報告書などを確定した後に解析レポートを最終化し、そのハッシュ値を記録することができる。

解析レポートパッケージの中の「測定機器データのリンク情報とそのハッシュ値」には、測定機器データパッケージの存在場所を示したリンク情報とそのハッシュ値を格納する。「測定機器データのリンク情報とそのハッシュ値」のファイルを作成するのではなく、測定機器データパッケージそのものを取り込んでもよい（図 5）。



破線がハッシュ演算の範囲。ハッシュ値はパッケージとは別に管理される。

図 5 解析レポートのパッケージ化（解析レポートパッケージ）

6. 今後の課題

5章で推奨する方法を実現するためには、以下のような技術の開発が期待される。

6.1. 各種測定機器データの標準化

5.1.にて示した、推奨する測定機器データの長期保存方法は、他のメーカーの測定機器のデータをインポートし、活用するのに適用できることが望ましい。このためにメーカーが互換性のあるデータフォーマットを提供する必要がある。現状では、3.3.4.で示されたものを含む標準データフォーマットがない測定機器が多い。

新しい技術と共に、新しいデータフォーマットが開発される。これらのデータフォーマットについては、機種に依存しないソフトウェアの提供、既存機器と互換のある形での出力、フォーマット定義の開示などが期待される。

6.2. データパッケージ技術の開発

5.にて示した、推奨する測定機器データ、解析レポートの長期保存方法を確実なものにするため

には、以下の技術の開発が期待される。

- a. 測定機器データ，解析レポートを自動的にパッケージ化する技術
- b. パッケージ化されたデータ，レポートを取り込み利用するための技術

7. 将来の展望

測定機器の高精度化と、1回の測定で取得する電子データの容量の増大が電子データの保存容量の増加をもたらしている。また、保存しているがほとんど活用されない電子データで、ストレージ（データ保存サーバ，データ管理システム等）の記録容量が占領されている事態が生じている。これらの結果、製薬業界では、電子データの保存に関するコストの増大が問題となりつつある。

製薬業界では、どの企業でも電子データの保存と管理について頭を悩ましていることから、電子データの保存と管理を受託するサービスのニーズがある（4.2参照）。更に、全てのメーカーの解析ソフトウェアの保存と管理を行い、どのメーカーの測定機器データも解析可能な環境を提供するサービスがあれば、製薬業界の多くの企業にとって有用であろう。

また、企業買収や関連する組織の提携などで、測定機器データの授受や、システム間でのデータ移行を行う必要も出てきている。古いバージョンのOSや解析ソフトウェアを仮想環境上で保持するサービスが可能であれば、企業が過去の測定機器データを新しいシステムに移行するなどの問題に頭を悩ませる必要もなくなる。

上記の対応は技術上可能と思われるが、そのようなサービスがビジネスとして提供されるためには、それを必要とする企業の数やその企業が支払うサービス対価がビジネスとして十分成り立つ規模である必要がある。ニーズの調査などを通じた実現可能性の調査を含めて今後の課題である。

8. 用語集

表 3 に、本ガイドンスで使用している用語の説明を示す。

表 3 用語の説明

用語	説明
アルゴリズム	明確に定義された有限個の規則の集まりであって有限回適用することにより問題を解くもの
インポート	他の測定機器で生成されたデータを利用できるように取り込むこと
エクスポート	生成されたデータを他の測定機器などで利用できるように書き出すこと
解析パラメーター	ピーク波形の処理方法などを設定した値。ベースラインの設定方法、ノイズ除去方法なども含む。
監査証跡	電子記録の作成、変更又は削除に関連する事象の過程の再構築を可能にする、安全でコンピュータで生成された、タイムスタンプ付き電子記録。オリジナルデータ生成のログ、解析の監査証跡、システムに関する監査証跡などがある。
QA	Quality Assurance, 品質を保証するための確認作業
サンプルスケジュール	インジェクションのスケジュール
ZIP	複数のファイルを一つのファイルとしてまとめて取り扱うアーカイブ用ファイル形式の一つ。基本的に zip を拡張子とするが、docx など拡張子が zip ではないファイルもある。
装置パラメーター	測定機器の動作条件などの設定値のほか、検出器の種類、溶媒の種類も含む。
タイムスタンプ	<ul style="list-style-type: none"> ・コンピュータが自動的に刻印する日時 (ERES 指針¹⁵) ・変更を検知できる情報を付与しその時刻以降に変更があったかどうかを証明する技術¹⁶
データインテグリティ	データの完全性 (completeness), 整合性 (consistency) 及び正確さ (accuracy)
データプロセス	データを処理する工程
電子署名	電磁的記録に対し、手書き署名又は捺印と同等のものとして行われる署名で、個人又は法人が作成、採用、確認、承認する一連の記号を電子化して構成したデータ (ERES 指針 ¹⁵)
パッケージ内ファイル構成情報	パッケージ内のファイル構成情報を記述した特別なファイル
ハッシュ値	ハッシュ関数により生成されたメッセージダイジェスト
PDF	Portable Document Format, ISO 32000 として規定された、文字、イメージ、グラフィックを含む文書の交換用ファイル形式
ビット	コンピュータが扱うデータの最小単位

¹⁵ 出典：薬食発 第 0401022 号平成 17 年 4 月 1 日 医薬品等の承認又は許可等に係る申請等における電磁的記録及び電子署名の利用について (ERES 指針)

¹⁶ タイムビジネス協議会 (<https://www.dekyo.or.jp/tbf/index.html>) タイムビジネス用語集

9. 改訂履歴

日付	版番号	改訂内容
2020.01.31	2.1	英訳作成に伴い、内容を見直し
2019.04.18	2.0	タイトルをガイドライン（案）からガイダンスに変更 募集したコメントの内容を反映
2018.11.28	1.0	

付録 1 AIA 形式データの例

```
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." ;// データファイルコメント
:sample_id = "1" ;// サンプル ID
:sample_name = "STD" ;// サンプル名
:sample_type = "Standard" ;// サンプルタイプ
:sample_injection_volume = 10.f;
:sample_amount = 1.f;
:detection_method_table_name = "";
:detection_method_comments = "This is comment." ;// データファイルコメント
```

```

:detection_method_name =
"C:¥¥Sample¥¥LC¥¥Demo_Method.met" ;
:detector_name = "検出器 A" ;
:detector_unit = "Volts" ;// 検出器の単位
:raw_data_table_name =
"C:¥¥Data¥¥Project1¥¥Demo_Data-001.dat" ;
:retention_unit = "Seconds" ;// 時間の単位
:peak_processing_results_table_name = "" ;
:peak_processing_results_comments = "This is comment." ;// データファイルコメント
: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 = "", "", "", "" ;
}
```

付録2 表2の詳細説明

本ガイダンス 表2に利用目的ごとの測定機器データの選択例を提供しているが、実際の運用においては各企業が状況に応じて判断していただきたい。

ここでは参考のため表2の区分の根拠を述べる。

表2（再掲） 利用目的ごとの測定機器データの選択の例

測定機器データ		再解析のみ	適合性 書面調査	GxP 規制 準拠
オリジナルデータ(機器固有フォーマットのデータ)		○	○	○
標準フォーマット (AIA) データ		○	○	○
派生データ	オリジナルデータを使って 解析された結果（ピーク面積等）など	—	○	○
	計算結果（濃度等）など	—	○	○
分析メタデータ	サンプルスケジュール	○	○	○
	装置パラメーター	○	○	○
	解析パラメーター	○	○	○
監査証跡メタデータ	オリジナルデータに関連するもの	—	○	○
	解析結果関連	—	—	○
	システム関連	—	—	—

○：格納する —：格納不要

再解析のみ

企業が測定機器データを長期間保存する目的の一つは、後日必要に応じて、再解析を行えるようにしておくためである。再解析に必要な要件を以下のように設定し、表2に反映した。

- ・ 再解析に使うソフトウェアに電子データが取り込めることが必須である。オリジナルデータが取り込めない可能性があるため、標準フォーマットデータも併せて保存しておく。
- ・ サンプルの特定や測定条件の情報源として分析メタデータを保存しておく。
- ・ パッケージにタイムスタンプを付与し、再解析をするまでに変更されていないことの証明とする。

適合性書面調査

新薬承認申請時に薬理試験、薬物動態試験・臨床薬物動態試験、品質試験などは、適合性書面調査において「申請資料の信頼性の基準（医薬品医療機器法施行規則第43条、第61条、第114条の22、第114条の42、第137条の25又は第137条の42）」に従って収集され、かつ作成されているものであるかどうかを調査される。上記基準に則って行われた試験の資料を保存することは当然のこと

であるが、長期保存方法や考え方には施設によって違いがある。ここでは以下の考えをもとに表 2 の要件を設定した。

- QA/QC などの社内監査部門の確認が終了したあと、オリジナルデータ、派生データ、分析メタデータ、オリジナルデータに関する監査証跡メタデータは、そのままの状態にパッケージに格納する。
- これに加えて長期保存後でも波形などの動的データの再現が可能となるよう、標準フォーマットデータもパッケージに格納する。
- パッケージにタイムスタンプを付与し、社内監査後の変更がないことの証明とする。

GxP 規制準拠

GxP 規制に則って試験資料を保存することは当然のことであるが、長期保存方法や考え方には施設によって違いがある。ここでは以下の考えをもとに表 2 の要件を設定した。

- QA/QC などの社内監査部門の確認が終了したあと、オリジナルデータ、派生データ、分析メタデータ、オリジナルデータと解析結果に関する監査証跡メタデータは、そのままの状態にパッケージに格納する。
- これに加えて長期保存後でも波形などの動的データの再現が可能となるよう、標準フォーマットデータもパッケージに格納する。
- パッケージにタイムスタンプを付与し、社内監査後の変更がないことの証明とする。

また、以下の理由により、ログインやログアウトなどのシステムの監査証跡はパッケージに含めないこととした。

- 本ガイダンスで提案しているパッケージは測定バッチ単位に作成するものであるが、システムの監査証跡は測定バッチ単位ではなくシステムの使用期間全体で管理するものであるため。
- システムの監査証跡はパッケージ化されたデータ以外に関する不適切な活動の検知に用いるものであるため。

(GxP 施設では、データ以外に関する不適切な活動は、QA/QC などの社内監査部門による定期的な監査等で確認している。)

以上

R & Dデータ保存委員会 ガイダンス作成メンバー（敬称略）

朝鳥 章	あすか製薬（株）
上原 小百合	アステラス製薬（株）
大野 治恵	旭化成ファーマ（株）
小川 泰弘	（株）LSIメディエンス
荻本 浩三	（株）島津製作所
木村 道弘	（公社）日本文書情報マネジメント協会
黒田 裕芳	（株）リコー
芝 清隆	（株）大塚製薬工場
鈴木 美代	生化学工業（株）
武田 幸雄	旭化成ファーマ（株）
平石 嘉昭	テルモ（株）
平城 里香	日本ウォーターズ（株）
細井 崇史	ジップインフォブリッジ（株）
宮地 直人	（有）ラング・エッジ
守野 智	エーザイ（株）
山崎 晃	協和キリン（株）
山田 宜昭	（株）日立ハイテクサイエンス

公益社団法人 日本文書情報マネジメント協会

〒100-0032 東京都千代田区岩本町 2-1-3 和光ビル7階

TEL 03-5821-7351 FAX 03-5821 7354

<http://www.jiima.or.jp>

法人番号 6010005003693