

測定機器データの長期保存 ガイドライン（案）



公益社団法人日本文書情報マネジメント協会
R&D データ保存研究会

2018年11月28日

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. | 従来の測定機器データの保存方法 | 4 |
| 3.3. | 再解析に関する背景, 懸念点 | 4 |
| 3.4. | 測定機器データの長期保存に関する懸念点 | 5 |
| 3.5. | データインテグリティに対する懸念点 | 6 |
| 3.6. | 測定機器データのフォーマット | 7 |
| 4. | 方針 | 8 |
| 4.1. | 解析結果に関する考え方 | 9 |
| 4.2. | 測定機器データの長期保存に関する考え方 | 9 |
| 4.3. | データインテグリティに関する考え方 | 10 |
| 5. | 推奨する測定機器データの長期保存方法 | 11 |
| 5.1. | 測定結果の長期保存方法 | 11 |
| 5.2. | 解析レポートの長期保存方法 | 13 |
| 6. | 今後の課題 | 13 |
| 6.1. | 各種測定機器データの標準化 | 14 |
| 6.2. | データパッケージ技術の開発 | 14 |
| 7. | 将来の展望 | 14 |
| 8. | 用語集 | 14 |
| 9. | 参考資料 | 16 |
| 10. | 改訂履歴 | 16 |
| 付録 | AIA 形式データの例 | 17 |

1. 目的

本ガイドラインの目的は、製薬業界の研究室で保存されている電子データを、信頼性を保ったまま長期間に渡り、容易に安心して保存管理する方法を提示することにある。なお、本ガイドラインで取り上げる電子データは、各種の測定機器から出力されるデータ（以下、「測定機器データ」という）であり、再解析を行うことを前提¹としている。

2. 適用範囲

本ガイドラインは、製薬業界における以下の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²で保存が義務付けられているデータ
- c. 基礎研究で得られたデータ

3. イントロダクション

IT 環境の急激な変化に伴い、組織が保存しなければならない電子記録の量も急速に増加してきている。また、紙記録とは異なり、電子データはシステム環境の更新で読み込めなくなることも多い。電子データを無思慮に保存していると、システムの更新やソフトウェアのバージョンアップなどにより、将来的に読めなくなることなどが危惧されている。特に製薬業界においては、電子データを信頼性が保たれ、再解析ができる状態で長期保存をしなければならないため、上記の課題はより重大である。

本ガイドラインを構築する背景や前提となった考え方を以下に示す。

3.1. 測定機器データの概要

測定機器、または測定機器に接続されたコンピュータには、生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などのデータ、並びにシステムのログが保存される。最近ではネットワークを介してデータベースに格納されることも多い。

これら測定機器等に保存されるデータは、次のように分類される。

- a. 生データ
測定機器から出力されたデータ。オリジナルデータと称されることもある。
AIA などの標準フォーマットに変換されたデータも生データに分類される。
- b. 派生データ
解析結果（ピーク面積等）、計算結果（濃度算出等）など

¹ 製薬業界では測定機器データが適切に解析されることが検証される必要がある。

² 患者の安全性や試験の信頼性を確保することを目的に公的機関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) : 「医薬品の非臨床試験の実施基準」

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、品質管理の基準」

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

c. メタデータ

測定条件，解析条件，サンプルスケジュール，電子署名，監査証跡など。このうち，測定条件，解析条件，サンプルスケジュールをこのガイドラインでは分析メタデータと称する。

3.2. 従来の測定機器データの保存方法

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は，電子データの状態のままではその保存性に確証が持てないため，プリントアウトして紙で保存することが一般的であった。この場合，筆記用具を限定³して第三者が署名・捺印をすれば，その記録の真正性が担保されると考えられる。しかしながら，次の問題があった。

- a. 再解析できない。
- b. 電子データの中に含まれているが，印刷されない情報（3D データ（断面しか印刷できない），解析条件，測定条件などのメタデータなど）は，すべて失われてしまう。また，印刷した監査証跡は信頼に耐えられないものとなる可能性がある。
- c. 紙の保管スペースが必要になり，社内だけで保管管理することが困難になる。

また，外部委託試験における施設間の電子データの引き渡しは，紙と比較して困難である。

- a. 紙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。
- b. 電子データを施設間で引き渡す場合には，真正性を保った状態で，どのような方法で引き渡すのか，引き渡されたデータをどのように保存するのかについて明確な基準がなく，各社が最適と考える方法をとらざるを得なかった。

3.3. 再解析に関する背景，懸念点

製薬業界では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，特に未報告の微小ピークを含めて調査する場合には過去のデータの再解析が必要となる。この時，もし適切に保存された測定機器データがあり，読み込める環境が用意できたとしても，他メーカーのソフトウェアで解析した場合，アルゴリズムの違いから，同一の結果が出るとは保証できない。

また，同一メーカーであっても，バージョンアップしたら，元の解析結果と同一の結果が得られるとは限らない。この原因には，次のものがある。

- a. 解析アルゴリズムの変更（解析精度の向上も含む）
- 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，小数点以下の計算結果に誤差を含むことはやむを得ない。
- 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かもしれない。

これらを防ぐには，解析が可能なコンピュータをソフトウェアまで含めて保存することが必要

³ 青色や黒色のボールペンなど書き換えたり消したりできない筆記用具。

となるが、現実的ではない。

なお、ここでは、天秤や pH メータの出力、確認試験における IR の結果は再解析の対象とはしない。

3.4. 測定機器データの長期保存に関する懸念点

再解析の必要性を考慮した場合、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。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データを移行できない恐れもある。

以下、代表的な長期保存方法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し、稼働できる状態を保持

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. データをバックアップできる状態の維持
- 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. データをバックアップできる状態の維持
- 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⁴、インタフェースカード⁵等追加のハードウェアが必要な場合の措置
- g. 仮想化の対象となるソフトウェアのライセンス形態⁶
- h. 完全に互換性がとれているかの検証が必要

⁴ 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）。

⁵ 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。

⁶ 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）

C) 現行システムから後継システム（または互換システム）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. メタデータとのリンクの喪失またはリンクの移行が困難である
- b. 移行元のシステムにデータとメタデータの一括出力機能が、移行先のシステムにその一括入力機能が必須である。
- c. 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。

本ガイドラインでは、「D) 標準フォーマットを介したデータ移行」を推奨する。

3.5. データインテグリティに対する懸念点

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、保管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、規制当局に提出する資料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。

当局の要求事項には次のようなものがある。

電子データの保存に対する当局の要求

FDA は “Questions and Answers on Current Good Manufacturing Practices, Good Guidance Practices, Level 2 Guidance - Records and Reports” というガイダンスの中で、次のような Q&A を掲載している。

- Q) ラボのコンピュータ化システムが生成する電子データに対して、Part 11 や cGMP はどの様に適用されるのか
- A) d. 製造過程や試験で使用された紙に印刷されたクロマトグラフィーデータ（クロマトデータ）は、cGMP の要件を満たしていない。
- e. 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データは、cGMP の要件に則って管理されなければならない。
- f. 電子データは、QC/QA 部門や FDA 査察官のレビューのために、管理され容易に利用できるようにしておかなければならない。

FDA はまた、次のような Warning Letter⁷を出している。

- a. 紙の試験結果は、全ての測定機器データを含んでいない。
- b. データに対してセキュリティ管理がされていない。

FDA のガイダンスでは、データインテグリティのために ALCOA を要求しており、EMA ではそれに加えて CCEA を要求している。日米欧の規制当局の期待に応じ、当局による査察（適合性書面調査など）に耐えるものとするためには、ALCOA-CCEA 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，普遍性）：記録の保存が永続的である |
| Available when needed（要時利用可能）：記録を必要な時に取り出せる |

また、MHRA は、2018年3月に発行した「‘GXP’ Data Integrity Guidance and Definitions」の 6.11.2 で、以下のように定めている。

- a. オリジナル記録とその真正なコピー（True copy）は、記録の完全性を保持しなければならない。
- b. 当該コピーの完全性を検証して記録する文書化された体系が確立されているなら、オリジナル記録の真正なコピー（例えば紙記録のスキャンしたもの）は、オリジナル記録の代わりに保管する場合がある。
- c. 組織は、オリジナル記録の破棄に伴うあらゆるリスクについて考慮すること。
- d. 真正なコピーは、記録の完全性を保つだけでなく、データの意味、監査証跡を含むべきである。

3.6. 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。あるメーカーのオリジナルフォーマットで保存し

⁷ 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

たデータを、そのフォーマットで他のメーカーの測定機器に移行することは難しい。

ここでは、よく使用されるクロマトデータを例として取り上げる。標準フォーマットととしては、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)

米国分析機器工業会 (Analytical Instrument Association - AIA) が定めた、クロマト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, FTIR, 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⁸。

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重要であるため (例えば、標準品を用いて検量線を作成し、濃度を算出する)、解析結果の絶対値を完全に再現する必要はないと考えられる。

また、データ形式が統一されていても、どのようなメタデータを取り込むかはメーカーによって違っており、必要とする項目にデータが入っていないケースや、データの形式 (単位、時刻など) が違うこともある。

4. 方針

本章では、5章にて推奨する方法の背景となる考え方を示す。表1に、3.3, 3.4及び3.5の懸念点・課題に対する方針のまとめを示す。

⁸ 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・課題 | 対応方針 |
|---|---|---------|
| 1 | 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。 | 4.1 に示す |
| 2 | 同一メーカーであっても、バージョンアップしたら、元の解析結果と同一の結果が得られるとは限らない。 | 4.1 に示す |
| 3 | メタデータとのリンクの喪失またはリンクの移行が困難である | 4.2 に示す |
| 4 | 移行元のシステムには生データとメタデータの一括出力機能が、移行先のシステムにはその一括入力機能が必須である | 4.2 に示す |
| 5 | 移行後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。 | 4.3 に示す |

4.1. 解析結果に関する考え方

3.3 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。生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、ソフトウェアに応じて、微妙な差が含まれることは許容せざるを得ない。いずれにしても誤差は小さいので、実質上、試験の目的に大きな影響は与えないと考えられる。

従って、解析結果は誤差を含むことを前提に、次のような観点から、データ保存に何を求めるのかを明確にすることにより、必要以上に要求しないことが望まれる。

- a. 再解析結果が元の解析結果とほぼ同等である必要があるのか。
計算機は必ず誤差を含んでおり、最小桁部分では結果が一致しない可能性があるが、これは許容される。
- 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差は許容されるのか。
- c. 前回と別の観点から見直すことができれば良いだけなのか。
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。

また、データ保存に対する要件は、結果再現、適合性書面調査対応、GxP 準拠など、利用目的や利用状況、およ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。

4.2. 測定機器データの長期保存に関する考え方

本ガイドラインは、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。このために、本ガイドラインは、以下の方法を提案する（図1）

- a. 測定機器から、生データや派生データのほか、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などのメタデータを一緒にパッケージ化してエクスポートする。
- b. このパッケージをストレージに保存し維持する。
 - c. このパッケージを、必要に応じてターゲットとなる測定機器、または、新たな解析環境にインポートして使う。

これにより、データの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を実現することが可能となる（7章参照）。

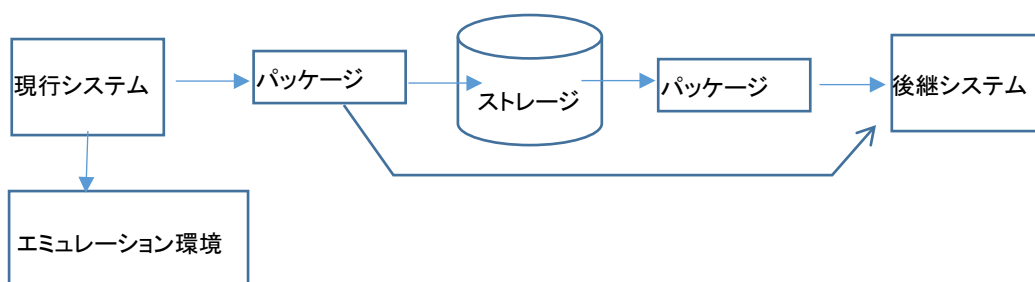


図 1 データのパッケージ化

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、生データ、派生データ、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）。エクスポートされるデータは、時間軸上のある時点のデータを切り取ったものとなる。

本ガイドラインでは、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。

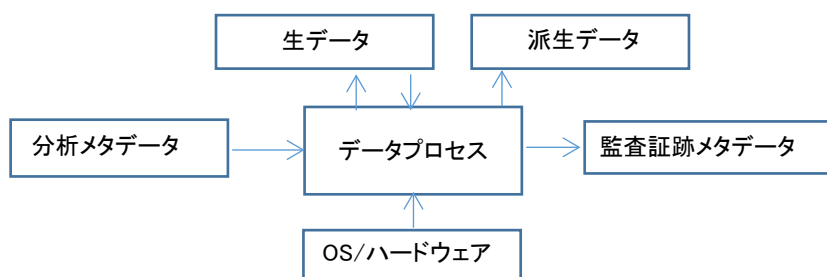


図 2 データインテグリティに係る要素

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、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の長期保存方法

再現性の確保に必要な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ドライン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。本ガイドラインでは、測定機器データと解析結果とで長期保存方法を別々に定義している。

なお、本ガイドラインでは例としてクロマトグラム（AIA フォーマット⁹）を取り上げている。

5.1. 測定結果の長期保存方法

測定機器から出力される標準フォーマットの AIA ファイルには、生データだけが格納され、監査証跡や解析パラメーターなどは含まれていない。付録に、AIA フォーマットデータの例を示す。

サンプルスケジュールは PDF ファイルとして出力できるが、AIA ファイルと関連付けられていない。

これらの関連付けには、複数のファイルをひとかたまりにして交換・保存することが有効である。そのための入れ物（パッケージ）としては、ZIP パッケージを利用することが推奨される¹⁰。

ZIP パッケージには、AIA ファイル、派生データやメタデータファイルに加えて、パッケージ内のファイル構成情報を記述した特別なファイル（META-INF）を格納する。ファイル構成情報を記述した特別なファイルは、ファイル名の命名規則により識別することになる。

パッケージ内の 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。

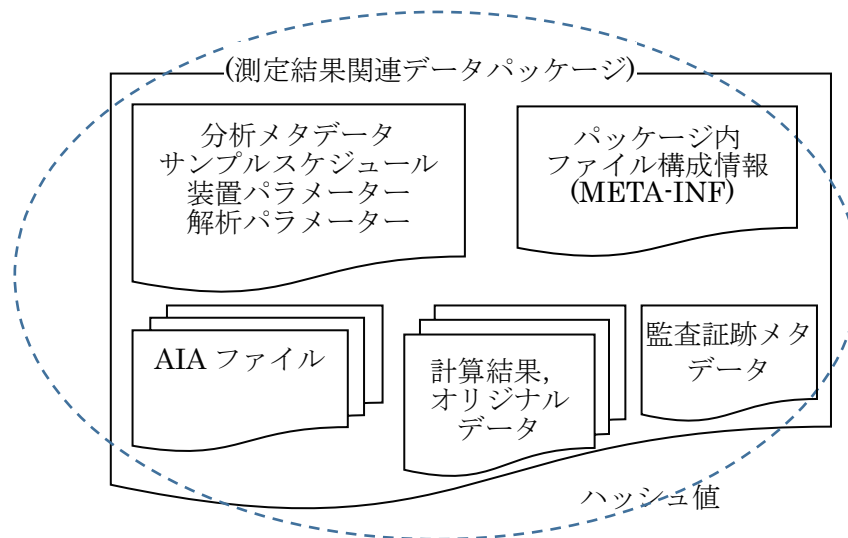
パッケージのイメージを図 3 に示す。このパッケージはあくまで測定機器データのもので、これらデータを用いて作成するレポートとは別のものである。パッケージに分析メタデータ及び監査証跡メタデータを含むことにより、測定機器データの真正性を確保できる。なお、クロマトデータでは、いくつかの目的の異なる試料を混在させて分析する場合もある。サンプルスケジュールに関連するデータを保存する際には、必要なものだけを抽出して保存するのではなく、他の目的のために測定したデータも含めて全て保存すべきである。

ハッシュ値の生成はタイムスタンプあるいは電子署名のどちらを用いても構わない。タイムスタンプを用いる場合、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方法などがある。

⁹ 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）

¹⁰ 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いうこともできる。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）。

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行うことで、オリジナルデータの誤消去や、上書き保存など行っ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 測定結果関連データのパッケージ化（測定結果関連データパッケージ）

パッケージに格納すべき測定結果関連データは、利用目的や規制要件によって異なる。表 2 に一例を示す。

表 2 利用目的ごとの測定結果関連データの選択の例

| 測定結果関連データ | | 結果再現 | 適合性書面調査 | GxP 準拠 |
|-----------|-----------------|------|---------|--------|
| 生データ | オリジナルデータ | ○ | ○ | ○ |
| | 標準化されたデータ (AIA) | ○ | ○ | ○ |
| 派生データ | 計算結果 | — | ○ | ○ |
| | オリジナル解析結果 | — | ○ | ○ |
| 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 | 生データ関連 | — | ○ | ○ |
| | 解析結果関連 | — | — | ○ |
| | システム関連 | — | — | — |

分析メタデータについては、個々のパッケージには含まずに、図 4 のようなメタ情報付与パッケージを構成して、外部から付与することも可能である。

なお、「測定関連データ」については、データそのものを格納するのではなく、データの存在場所を示したリンク情報を格納することで、他の目的ですでにパッケージ化されているデータを参照したり、保存データの容量を削減したりすることができる。

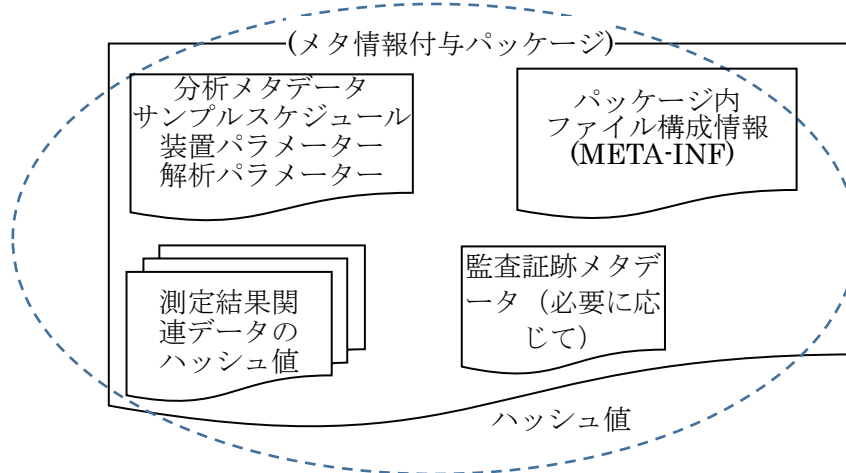


図 4 測定結果関連データへのメタ情報の付与（メタ情報付与パッケージ）

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の変更や、後日異なる角度からの再解析を行ったり、別の解析結果との比較を行ったりするなど、修正することもある。しかし、試験報告書などを完成した時点で一旦それを最終化し、ハッシュ値を記録することで真正性を確保する。ここに同梱する測定結果関連データパッケージおよびそのハッシュ値は、オリジナルのパッケージそのものを取り込んでも良いが、データの保存場所を示した情報またはハッシュ値でもよい。（図 5）

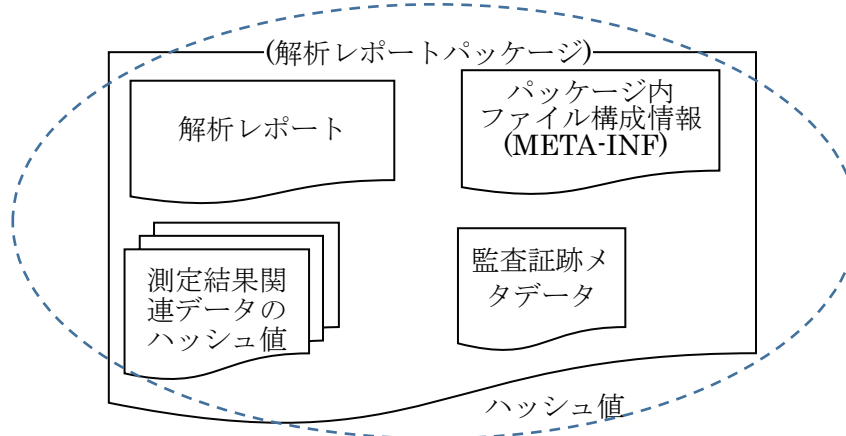


図 5 解析レポートのパッケージ化（解析レポートパッケージ）

6. 今後の課題

5章で推奨する方法を実現するためには、その方法を実装するための技術が必要になる。本ガイドラインは、各種測定機器データのフォーマットとデータパッケージ技術の開発を推奨する。

6.1. 各種測定機器データの標準化

5章で推奨する方法は、他のメーカーの測定機器のデータであってもインポートして活用できることが望ましい。そのためにはメーカー互換性のあるデータフォーマットの提供が必要になる。現状は、3.6項で示したデータフォーマットの対象となる機器は限定されており、標準的なデータフォーマットのない測定機器も多い。

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、機種非依存のリーダ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章で推奨する方法を確実なものにするためには、自動的に測定機器データをパッケージ化する技術の開発が期待される。

開発が期待される技術の概要としては、以下の通りである。

- a. データを自動的にパッケージする技術
- b. パッケージ化されたデータ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回で取得するデータの容量の増大が保存容量の増加をもたらしている。また、ほとんど活用されない保存のためだけのデータでサーバーの記録容量が大幅に占領されてしまう事態となってきた。これらの結果として生ずるコストの増大が問題となりつつある。

製薬業界では、どの会社でも電子データの保存と管理については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節参照)。更に、全てのサプライヤの解析ソフトウェアの保存と管理を行い、どのようなサプライヤのデータも解析可能な環境を提供する施設があれば、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間の提携などでのデータの授受や、システム間でのデータを移行させる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ことが可能であれば、利用企業が過去の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はなくなる。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成り立つためには、それを必要とする組織の数とその組織の提供可能なコストが十分である必要がある。この構想については、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、本ガイドライン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 |
|------------|---|
| IR | Infrared spectroscopy の略称。赤外分光法 |
| 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 |
| AIA ファイル | ASTM で定義しているクロマトデータのやり取りを効果的に行うためのフォーマット |
| エクスポート | 生成されたデータを他の測定機器で利用できるように書き出すこと |
| オリジナル分析結果 | 測定機器から得られる結果で、必要な情報はほぼ全て格納されているが、装置メーカーや型式、データ処理プログラムのバージョンによってもデータ形式などは異なる場合がある。 |
| 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、変更また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 生データ生成のログ、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 |
| 計算結果 | クロマトデータの波形から、ピーク位置、ピーク面積などを求めた結果のファイル |
| 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 | ZIP を拡張子とする、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 注記 docx など拡張子が ZIP とは限らない |
| 真正性 | 文書情報を取り扱う組織が組織活動の目的、権限及び業務によって作成された文書情報を作成時から運用維持されていることが説明できる状態 |
| 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 | 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 |
| データインテグリティ | データの完全性(completeness)、整合性(consistency)および正確さ(accuracy) |
| データプロセス | データを処理する工程 |
| 電子署名 | 文書情報に作成者及び/又は承認者の承認情報を与える仕組み |
| パッケージ内構成情報 | 測定機器データを長期間保存するためのパッケージが、どのような情報を含んでいるかを明示したもの。 |
| 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 |

9. 参考資料

1. “Questions and Answers on Current Good Manufacturing Practices, Good Guidance Practices, Level 2 Guidance - Records and Reports” FDA,
(<https://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm124787.htm>)
2. ‘GXP’ Data Integrity Guidance and Definitions, MHRA
(https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/687246/MHRA_GxP_data_integrity_guide_March_edited_Final.pdf)
3. 「タイムスタンプ長期保証ガイドライン」タイムビジネス推進協議会，平成 18 年 4 月

10. 改訂履歴

| 日付 | 版番号 | 改訂内容 |
|------------|-----|------|
| 2018.11.28 | 1.0 | |

付録 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 = "", "", "", "" ;
}
```

EOF