資料②

適切な医薬品等の安全性評価活動に資する電子カルテシステムの診療情報等を情報源とする医療情報データベースの信頼性確保に関するガイダンス（試案）

研究代表者　横井 英人

厚生労働行政推進調査事業費補助金
医薬品・医療機器等レギュラトリーサイエンス政策研究事業
23KC2001

令和 8 年 5 月

資料②

1. 目的及び基本方針

本ガイダンスは、厚生労働省通知「医薬品の製造販売後データベース調査における信頼性担保に関する留意点について」を踏まえ、医療情報データベース（以下「医療情報 DB」という）の品質管理における医療機関、医療情報データベース事業者（以下「DB 事業者」という）間の連携や電子カルテを管理するために重要な点をとりまとめる。同通知には、製造販売業者等が製造販売後データベース調査を行う上で、再審査等の申請資料の信頼性を担保する観点から留意する必要がある事項が記載されており、その別添には、製造販売業者等が確認すべき DB 事業者の医療情報 DB に係る手順書として DB 事業者が整備するべき事項について言及されている。そして、同手順書には医療情報 DB を利用した医薬品の再審査等の申請を行う際に求めるべきデータベースの要求品質が DB 事業者の観点で整理されている。医療情報 DB の品質は DB 事業者と医療情報 DB の情報源である医療機関との円滑な連携の上で担保されると考えられるが、その観点から医療情報 DB の品質管理について言及した文書は未だ存在しない。そのため、本ガイダンスは同通知が示す DB 事業者が医療情報 DB の信頼性を担保する上で実施すべき事項における医療機関との連携に焦点を当て、そのポイントを整理する。具体的には、DB 事業者と医療機関等との円滑な連携が確保された上で、医療情報データの二次利用を見据えた電子カルテ等の維持管理及び医療情報 DB の品質管理を適正に実行するための指針を示す。また、その際に、医療情報 DB の品質管理の国際整合性の確保を鑑みて、他国の医療情報 DB の特徴、データ集積フロー、品質保証の取り組み等を参考にしたうえで、我が国における医薬品等の安全性評価活動に資する医療情報データベースのデータ品質を向上するための要点を整理する。

2. 適用範囲

本ガイダンスが対象としている事業者の範囲は、①医療情報 DB 事業者（DB 保持者、管理者等)、②医療情報 DB の情報源であり、患者等に直接医療を提供する施設（医療機関等）である。

3. 医療情報データベースの品質管理の考え方と課題

医薬品の安全性評価において医療情報 DB を用いた有効な分析を実施するためには、①医薬品の安全性評価に必要なデータ項目があり、妥当な分析の結論を得るために必要な症例数・代表性（悉皆性があれば最も望ましい）があること、②登録されているデータが医療実態を正確に反映したデータであることが求められる。また、その上で、データの安全で継続的な収集を可能とするため、③収集したデータがセキュアな状態で適切に管理され、必要なユーザーがデータにアクセス可能であることが求められる。

上記の中で、①及び③は国内外にて多くの医療情報 DB にて達成されており、データベース構築の際に医療情報 DB 事業者が検討する事項である。一方で、②を実現するためには医療情報 DB 事業者のみならず、医療機関における電子カルテシステム等への診療データのインプットとそのセキュアな管理、および医療情報 DB への診療データの ETL(Extract/Transfer/Load)への的確な対応が重要であると考えられ、以下の品質が求められる。

1. データ収集時において医療機関にて電子カルテシステム等に入力した診療情報は正確である

2. ユーザーにデータが届くまでのデータの完全性・一貫性が担保されている（医療情報 DB の構築・管理運営の過程においてデータの完全性と一貫性が成立している）
3. データの利用用途（医薬品の安全性評価）或いはデータの二次利用を意識したデータ収集が行われている

本文書は、上記について、現在、我が国で製造販売後の医薬品の安全性を評価することを目的に構築された医療情報 DB である MID-NET®や海外での同様の医療情報データベースの品質管理の事例を参考に、電子カルテシステムの診療情報等を情報源とする医薬品等の安全性評価活動に資する医療情報データベースの品質管理における落とし穴を紹介する。

3.1 MID-NET®にて実施しているデータの品質管理

MID-NET®の運営・管理を行っている医薬品医療機器総合機構（以下、「PMDA」という）では、MID-NET®システムのデータ全体を対象に、データの移行状況を日次モニタリング、統合データソース*のデータ格納状況を月次モニタリング、及び統合データソースに格納されたデータの元データとの一致性を定期的にモニタリングしている。また、標準コードの対応表を継続的に（年 2 回程度）管理することでデータの品質管理を行っている。その中で認められたデータの件数及び内容についての課題を一部紹介する。

3.2 MID-NET®の品質管理で認められた課題

MID-NET®で利用可能なデータのうちレセプト、DPC 由来のデータについては、医療機関において、国が定めた形式でデータを作成しているため、すべての医療機関でデータ形式が統一されている。一方、電子カルテ由来のデータについては、医療機関ごとに形式が異なるため、データ収集にあたり標準化が必要となる。このため、統合データソースにデータを移行する際に、データ形式の変換を行う必要があり、課題が発生しやすく、電子カルテ由来のデータで元データとデータの件数や内容が一致しないケースが多く認められる。

データの件数が一致しないケースは、送信されるべきデータが送信されていない、不要なデータが送信されているといった様々な状況により発生する。

＜具体例＞

- ✓ 電子カルテに持参薬処方として登録されているデータがデータを移行するプログラム（以下、「データ移行プログラム」という）で移行対象外と設定されていたため、送信されるべきデータが送信されていなかった

- ✓ 電子カルテに指示のために登録しているデータ（持参薬報告、中止指示等）がデータ移行プログラムで送信対象と設定されていたため、不要なデータが送信されていた
- ✓ データ移行プログラムの不備で、データが正しい順番で移行されず、必要なデータが削除されていた

また、件数だけでなく内容についても一致しないケースがある。こららは、データ解釈の誤り、データ取得先の誤り、標準コードの設定の誤りなど、様々な原因により発生する。

<具体例>

- ✓ 処方方法（1 日量処方、1 回量処方、全量処方等）の違いによる処方量が示している値をデータ移行プログラムで正しく解釈できておらず正し値が設定されていなかった（通常、電子カルテで登録されている処方量は、処方方法（1 日量処方、1 回量処方、全量処方）によって値の解釈が異なる）
- ✓ 入院中に外来受診した患者の診療科が、データ移行プログラムのデータ取得先の設定誤りのため、入院診療科の情報が外来受診診療科として移行されていた
- ✓ 電子カルテで検体検査の結果値を検査項目によって異なるフィールドで登録していたため、データ移行プログラムが正しいフィールドからデータを移行しておらず、電子カルテと異なる値が移行されていた
- ✓ 検体検査の測定法が変更されたが、医療機関において同一のローカルコードを継続利用していたため、正しい標準コードが設定できなかった

上記確認された課題は、医療機関からデータを移行する際のデータ移行プログラムの設定誤り、データ移行プログラム不備によるものが多いため、医療情報 DB の構築段階で品質管理を実施していれば問題ないと思われがちだが、病院情報システムは、定期的にシステム更改が行われるため、これまで問題なく送信されていたデータが、病院情報システムの更改のタイミングで、正しく移行されなくなるという事象が発生する。また、システム更改だけではなく、部門システムの変更や医療機関における運用変更も正確なデータ移行に影響を与える場合があり、データの不一致は医療情報 DB の運用開始後も認められる。

3.3 DB 事業者における医療情報 DB の品質管理

前項にて、MID-NET®を運営・管理する上で PMDA が確認したデータ品質に関する課題を紹介したが、多くの課題は、電子カルテから医療情報 DB にデータを移行する部分で発生しており、特に、医療機関の電子カルテの使い方に起因するケースが多い。医療機関で発生するデータは、電子カルテの取扱事業者の製品ごとにデータの持ち方や管理方法が異なり、医療機関の運用にあわせて電子カルテ自体がカスタム化されることもある。医療機関から医療情報 DB へのデータ移行プログラムは、電子カルテ取扱事業者が標準パッケージとして準備している場合も多いが、電子カルテをカスタム化している場合、カスタム化していなくても医療機関の運用にあわせた独自の使い方をしている場合があるため、医療情報 DB へのデ

ータ移行における品質の確認は、電子カルテ取扱事業者の標準パッケージでの検証（バリデーション）のみでは不十分であり、医療機関ごとに実データを用いた確認が必要であると考える。また、医療情報データベースを構築した後も、医療機関の運用変更などが発生するため、医療情報 DB 構築時だけでなく、データ移行プログラムに影響を及ぼす可能性がある変更が医療機関側で実施されていないかを継続的に確認する必要がある。加えて、電子カルテや部門システムは定期的にリプレースや更改が発生するため、電子カルテのリプレース等のたびに、医療情報 DB 構築段階と同じような実データを用いた検証・確認を実施する必要がある。

4. 医療情報 DB の品質管理における留意事項

4 章では、上記課題を踏まえ、医薬品等の安全性を評価する目的で医療情報を収集・活用する際の、医療情報 DB の品質管理のポイントを項目別に示す。

4.1 医療機関における電子カルテシステムでの診療データの管理

4.1.1. 電子カルテへのデータ入力

医薬品の安全性評価に使用される診療情報は電子カルテデータ、レセプトデータ、DPC データが主なデータソースである。医薬品安全性評価においては、特に当該医薬品の使用に関する情報（投与日や投与量など)、臨床検査に関する情報（測定結果など）および疾患や症状に関する情報などの品質を担保する必要がある。また、複数の電子カルテ取扱事業者の製品が医療機関にて利用されている中で、データ項目の標準化及び入力ルールの統一が重要である。

本邦においては、医薬品副作用情報を患者プロファイル等に任意で記載する形となっている上、これらの情報の薬剤標準マスターとのマッピング可否は電子カルテ取扱事業者ごとに異なっており、統一がなされていない。そのため、医薬品副作用情報に関する構造化データを取得することは困難である。また、医薬品の安全性評価に使用される医薬品、臨床検査、および疾患や症状に関する情報の電子カルテ入力時点での標準コードへのコーディングについては、医薬品（YJ コード等)、傷病名（ICD-10）については、多くの電子カルテ取扱事業者の製品で対応ができるものの、臨床検査（JLAC10）については、コーディングが出来ない、あるいは外付けの変換マスターが必要な製品が多い。

また、電子カルテにおける傷病名については診療報酬請求のためにレセプト記載に用いる傷病名（レセプト病名）が記録されていることがある。当該病名は ICD-10 とのマッピングはできているものの、医学的な傷病名を別途付与される場合も多く、これらの病名は、プロブレムリスト・既往歴として任意で管理する運用がされている。傷病名の開始日・終了日については、レセプト病名の開始日（初診日)・終了日（転帰確定日）が記録されている。レセプト病名の開始日・終了日とは異なる本来の開始日（発症日)・本来の終了日（治癒日等）が別途付される場合、上記同様プロブレムリスト・既往歴として任意で管理できる形となっている。本来の開始日・本来の終了日のデータの持ち方は電子カルテ取扱事業者ごとに異なっており、統一がなされていない。

このような現状から、特に傷病名については、医学的に妥当な傷病名を特定するには、“当該患者における医薬品の使用状況や検査結果等の情報を加味した形での傷病名の特定”が必要になる。すなわち、ある傷病名を医薬品の副作用として利用する場合には、その傷病名およびそこに付与された標準コードの医学的な妥当性の見極めが重要になる。この点は、医療機関の診療情報を情報源とする医療情報 DB を利用した医薬品の安全性評価における必須の課題の一つであることから、アウトカムに傷病名を利用する場合には、アウトカムのバリデーションが重要になる。

4.1.2. 入力されたデータの品質保証

医療情報 DB にてユーザーが利用するデータの正確性は、診療時に入力されたデータの正確性に依存しているが、4.1.1 に記載したように、レセプトデータと整合するように電子カルテにデータが登録されているケースがあるなど、データの正確性よりも整合性が優先されるケースが散見される。例えば、検査の段階において疑い病名は記載しないことが多いため、レセプトデータでは、検査をする過程で付与される病名と、治療が必要となった際に付与される病名が異なる場合があり、電子カルテにおいてもレセプトデータと同様に病名に疑いレベルを示すことができないため、その区別ができないことも課題である。加えて、複数の適応症をもつ医薬品の場合には、医薬品とその医薬品の使用目的となった傷病名とを直接紐づける情報を、電子カルテで示すことができないため、傷病名と医薬品を用いただけでは、正確にアウトカムが特定できないケースも存在する。

傷病名の正確性については海外のデータベースを見習うところがある。医療情報の収集・蓄積過程や使用されている電子カルテが我が国とは異なるため直接比較はできないものの、欧米においては、医師とメディカルコーダーの連携、ソフトウェアを活用したデータの整合性チェック、及び監査を実施することで入力される傷病名の品質を向上させている。英国では、primary care において general practitioner (GP)が自ら診療情報の正確性を確認し、GP によっては更に EHR に搭載されている compliance software にて統計データに基づくチェックをした後、医療コーダーによる不整合のチェックを行い、正確性が確認される。また、医師のコーディングが監査対象となっており、年に一度程度、データ品質のトレーニングやドキュメンテーションのチェックが実施されることによりコーディングの正確性向上の意識づけもされている。米国においては、メディカルコーダーがコーディングする際に電子カルテデータの正確性を確認する仕組みがあり、監査による品質保証もされている。

これらの例を参考に、コンテキスト情報を活用しながら、正確なコーディングをどう実現するかを検討することは重要であると考える。日本において多数の電子カルテ取扱事業者の製品が利用されている現状、保険請求制度・医療現場における文化の違いを鑑みると英米のようなソフトウェアの普及等は困難であると考えられるが、医療情報 DB 事業者は医療情報従事者等により入力されたデータの正確性を確認する仕組みを導入するなど入力されたデータの品質が保たれるよう、医療機関と連携しルール・ポリシー策定を進めることが望ましい。

4.2. 医療情報 DB へのデータ移行時のデータの完全性・一貫性

医療機関にて電子カルテ等に入力された医療情報は、通常、複数のデータ集約・統合過程を経て最終的に医療情報 DB に移行される。その過程において、データの完全性と一貫性が担保されている必要がある。本邦においても例えば MID-NET®では、MID-NET®を運営・管理している PMDA が医療機関と協力しながら院内システム、標準ストレージ、統合データソース間とデータの移行の各ステップにおいてデータの確認を行い、見つかった課題に関して対応を行っている。

一方、英国（CPRD)、EU（DARWIN EU)、米国（Sentinel Initiative）においては、ソフトウェアを活用したアルゴリズムによるチェックが主である。英国はデータベースの孤立レコード（他データと紐づいていないデータ）を特定することで、データ登録の不整合等の検知、及びデータ項目ごとの値の「許容範囲」を設定した上で、異常値を検知するアルゴリズムを利用している。米国では、data provider（保険会社等）からデータを収集する際に、完全性、データ項目、テーブル間での一貫性や、過去抽出データとの整合性を評価するアルゴリズムを利用している。研究プロジェクト単位で情報収集を行うため、必要に応じて FDA がデータのバリデーションを data provider と実施することでデータ移行時に発生する問題に対処が可能な仕組みになっている。

先述の通り、上記 2 か国の電子カルテシステムや医療情報 DB の仕組みは我が国のそれと異なるため、品質管理の有効性を比較することはできないが、医療情報 DB へのデータ移行について、医療情報 DB 事業者は、データの完全性と一貫性を担保することは重要であると考える。3.3 に記載の通り、多数の電子カルテ取扱事業者の製品が利用されている現状、医療機関の運用にあわせ電子カルテのカスタム化が医療機関ごとになされている現状を鑑みると、本邦においては、実データを用いた品質保証のプロセスを構築する必要があると考える。しかしながら、医療情報 DB に蓄積されるすべてのデータの品質を保証することは、リソースやコストの観点からも現実的ではないため、品質保証のプロセスを実行することにより解消できる問題とデータベースユーザーが求めるデータ品質とを照らし合わせ、医療情報 DB の仕組みに合った品質保証を実施することが重要である。また、電子カルテ等に入力された医療情報についての完全性および一貫性の品質保証において生成 AI の利用によって従来の人による目視確認といった作業負荷が改善され、その作業効率の向上が図られることが期待される。

4.3. 二次利用を意識したデータ収集

データの正確性は、入力者のデータ入力に関する考え方に左右されることがあるため、データの正確性を向上するためには、医療機関においてデータの利用用途を認識・意識してデータを入力すること、二次利用を意識して医療情報 DB にデータを移行することが重要である。

他国がどのようにこの点に対応しているか見渡すと、例えば米国では、研究課題を設定した上で、データ収集を行っているため、データの二次利用ではなく、一次利用に近い形式をとっている。具体的には、米国の Sentinel Initiative を運営する FDA が、研究プロジェクト毎に Data Provider（データベースにデータを提

供する医療機関や保険会社）と連携をし、プロジェクトに必要なデータを収集する。必要に応じて医療機関への問い合わせや、バリデーションを行い、データの意義や正確性を確認した上でプロジェクトを遂行する。個別医療機関レベルでのデータ二次利用の意識づけは限定されるものの、プロジェクトに必要なデータをカスタムで収集する仕組みが一部存在する。欧州の DARWIN EU でも同様に分散型データ管理を行っており、研究プロジェクト毎にデータの収集方法、関与する Data Partner が異なり、目的に即した調査設計及びデータセットの構築が可能である。また、英国においては EHR の構造化されたデータを利用しており、研究用途でデータ利用されることを提示している。そのため、医師によって程度は異なるものの、一部意識づけが進んでいる。

一方、日本では、医療情報 DB を構築、データを蓄積した上で研究課題に合ったデータを抽出するという形式の医療情報 DB がほとんどである。医療情報 DB の構築、データ蓄積段階では、具体的なデータの利用用途は未確定のため、医療機関が日常診療において、二次利用を意識してデータを入力することは難しく、これらのデータを二次利用のために医療機関自らがもれなく標準化することも現状では難しい。

このため、本邦においては、医療情報 DB 事業者と医療機関が協力し、医療機関に二次利用を意識したデータの入力やデータを収集するための体制を整備することが重要となる。他方で、すべての医療機関に二次利用を意識したデータの入力、収集を義務化することは現実的ではないため、医薬品の安全性を評価することが目的の医療情報 DB では、対応可能な医療機関をデータ収集先に選定する必要がある。また、データの利用用途（医薬品の安全性評価）によっては、データの悉皆性よりもアウトカムバリデーションを実施しているなどの評価対象となるデータの品質を優先して、利用する医療情報 DB を選択するといった判断をすることも重要であると考える。

さらには、生成 AI（Generative Artificial Intelligence）の利用による電子カルテの入力データのメタデータの付与も診療情報の二次利用を意識したデータの品質管理の方策として有用と考えられる。例えば、電子カルテシステムには 1 患者につき複数の傷病名が記録されている。それらの傷病名は既往歴や合併症、あるいは検査実施のための疑い病名など様々な背景情報を有するが、その背景情報について生成 AI を用いて各傷病名のメタデータを作成し、電子カルテシステム内で共有することは、傷病名に係るデータを二次利用する場合に極めて有効な情報と成り得ることが期待される。

医療機関におけるデータの標準化については、医療情報のコーディングポリシーの統一と、データの二次利用のための標準化に関する教育が重要である。医療情報 DB 事業者は医療機関と連携し、上記ができる体制の整備を支援することで入力データの正確性向上が期待できる。

5. 統括

本ガイダンスは、「医薬品の製造販売後データベース調査における信頼性担保に関する留意点について」の目的を達成するため、DB 事業者と医療機関の連携や電子カルテを管理するために重要な点を、他国の医療情報 DB 等を参考にしたうえで整理した。品質の向上に向けた課題は DB 事業者のみの努力だけでは解決はしなく、事業者が医療機関の特性を理解し、医療機関との連携方法を考慮すべきである。特に、データの二次利用に関する意識づけについては充分に医療機関に説明することが重要であり、データの収集対象とした医療機関の特性を理解してデータベースを構築・活用するべきである。医療機関は本来、データの利用用途については考慮しないことが多いため、データ登録の際に重視すべきことを伝えることでデータの品質向上が期待される。

資料②

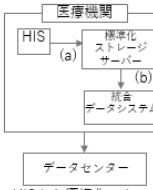
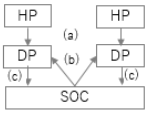
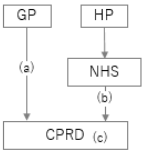
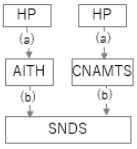
■　参考資料

医療情報DBの比較 ①

項目	MID-NET (JPN)	Sentinel Initiative (USA)	CPRD (UK)	SNDS (FRA)
利用者	• PMDA • 製薬企業 • 研究者など	• FDA ※製薬企業等の依頼により当局が有料で分析可能	• 製薬企業 • 研究者 • 規制当局　など	• 製薬企業 • 研究者など
DBに格納されている医療データの種類	• 患者情報、来院等情報（外来、入院、退院）、傷病情報（退院サマリ、病名オーダ） • 放射線検査、生理検査 • 検体検査、細菌検査 • 処方情報（オーダ・実施）、注射情報（オーダ・実施）など	• 管理データ(基本情報、診断、処方など) • 臨床データ • 死因データ • 補助データ(施設IDなど) • ワクチンデータ • 入院データ(調剤、輸血など) • 患者報告測定データ • 母子系統データ	• 人口動態データ • 診断と症状 • 臨床検査 • 薬物曝露 • 予防接種歴 • セカンダリケア • 診断コードが添付された処方情報	• 健康保険データ • 病院データ • 死因データ • 障害に関するデータ • 健康保険組織からのデータ
データソース	• 電子カルテ • レセプトデータ • DPCデータ	• レセプトデータ • 電子カルテ（限定的）	• GPが入力する患者データ（Vision®或いは EMIS®のEHRデータ） • NHS等のセカンダリケアデータ（Hospital Episode Statistics (HES)データ）	主に以下の２つ • AITH (入院情報技術庁) • CNAMTS (国民健康法権基金)
規模	23機関 （2018年4月1日現在）	20～50機関	約6,500のGP施設のうち、約2,200施設(全体の34%)	ATIH病院約1,900施設（内科、外科の例）

1

医療情報DBの比較 ②

項目	MID-NET (JPN)	Sentinel Initiative (USA)	CPRD (UK)	SNDS (FRA)
データ集積フロー概要	 a. HISから標準化ストレージサーバーにコード化されたデータが移行される b. 標準化ストレージサーバーから統合データシステムへデータが移行される HIS: 病院情報システム	 a. DPは各医療機関のclaimsデータを収集し、自社形式で整理した後、Sentinel Common Data Model（SCDM）形式に変換する b. SOCが各DP（保険会社等）に必要なデータについてクエリを行う c. DPがクエリ結果としてSCDM形式のデータをSOCに返す HP: 医療機関 DP: データパートナー SOC: Sentinel Operations Center	 a. Primary Care: CPRDは参加同意を得たGPから毎月特定形式のEHRデータを収集する b. Secondary Care等: NHSは患者の NHS番号、生年月日、郵便番号、性別を仮名化し、CPRDに送付する c. CPRDでaとbの紐づけを行う（縦断データの構築）	 a. 医療機関はハッシュ関数を用いて患者データの仮名化を行い、各国當DBにデータ送信する b. 各DBにて再度、ハッシュ関数を用いた仮名化を行いSNDSにデータを送信する

5

資料②

医療情報DBの比較 ③

項目		MID-NET (JPN)	Sentinel Initiative (USA)	CPRD (UK)	SNDS (FRA)
	データ収集目的（再掲）	医薬品安全対策の質の向上のため	医薬品安全対策の質の向上のため	公衆衛生と臨床研究をサポートするためのデータ収集	公衆衛生と臨床研究をサポートするためのデータ収集
品質保証の取り組み	① Practitioner recording	**実施者:** 医療機関 • 厚労省が発出している医療情報システムの安全管理に関するガイドラインに沿って、医療機関にて独自にDBの情報源となる医療情報システムとしての品質管理を行う	**実施者:** 医療機関・保険会社 • 医療機関がmedical data entry specialist等と連携し、clinical validation等の監査プロセスによりClaimの妥当性を検証した上で構造化したデータをDPとなる保険会社等に発信する	**実施者:** GP • GPがICD-10やSNOMEDを利用しコーディングを行い、特定ソフトウェアにてEHRを作成しCPRDに発信する際にQRMシステムや専用ソフトウェア等を利用し独自に品質保証を行う	**実施者:** 医療機関・保険機関 • 病院で医師主導の専属チームがICD-10等を利用し電子カルテのデータのコーディングを行い、各国営データベースに発信するが、入力内容は保険機関の監査により信頼性を確認する
	② Data Extract Transform Load	**実施者:** 協力医療機関、PMDA • 統合データベースへのデータの抽出と転送に関して協力医療機関の共通の管理手順書に基づいた対応が実施され、変換については、九州大のガバナンスセンターにおいて薬剤および臨床検査に係る変換テーブルのコード管理が行われている。 • システム導入時の品質管理に加え、データ・システムの日常的なモニタリング及び協力医療機関を年1～2回訪問し、データの信頼性を確認を行う • HIS、標準ストレージ、統合データソース間でデータの比較を行い、見つかった課題に関して対応を検討する	**実施者:** SOC • DPにて独自のデータ形式からSentinel形式に変換する際等、品質管理を行うことが多い（任意） • Sentinel形式のデータをSOCにて収集する際に以下の品質保証を実施する 1. 完全性やデータ項目 2. テーブル間での一貫や時系列の整合性 3. 前回抽出データとの比較	**実施者:** CPRD • GP送付データの項目、構成要素、長さ、形式を確認し整合性をチェックする • データベースの孤立レコード（他データと紐づいていないデータ）を特定することで、データ登録の不整合等を検知する • データ内容はデータ項目ごとの値の「許容範囲」を設定し、異常値を検知する	**実施者:** AITH • 医療機関から受領したデータの品質管理を行うツールを複数利用し、データの整合性やコーディングのエラーを検知し、課題のある医療機関を特定する
	③ Maintenance of Database	• 監視により信頼性・完全性及び保存情報の機密性を確保する	• Sentinelではデータセットを保持せず、データ利用の際にプロジェクトごと収集したデータ品質を確認する	• NHSデータセキュリティおよび保護ツールキットの年次評価を実施する	• トレーサビリティ、認証状況、データの整合性を監視により確保する

医療情報DBの比較 ④

項目	MID-NET (JPN)	Sentinel Initiative (USA)	CPRD (UK)	SNDS (FRA)
データクリーニング	標準化データ出力と統合データソースの格納前に実施される	DPにて簡易的に実施している場合もあるが、Sentinelでは実施していない	内部プロセスとして、重複データの排除などのダブルチェックのみ実施している	病院ではソフトウェアを用いて実施されることはあるが、データ発信後は実施しない
個人情報保護	【関連法規】 PMDA法 【データ加工】 患者氏名、ID、住所は削除 【同意取得】 医療機関からPMDAに、PMDAから医療機関に個人情報を提供することに、本人の同意は不要である（オプトアウトはあり）	【関連法規】 ・HIPPA ・FISMA (Federal Information Security Modernization Act) 【データ加工】 非識別化情報またはリミテッドデータを利用し、個人を特定できる情報を含んだデータは受信しない 【同意取得】 患者とDP間で行われるが、非識別化情報またはリミテッドデータを利用するため同意なくデータの使用・開示が可能である	【関連法規】 National Data Guardian 【データ加工】 CPRDは患者氏名、住所、NHS番号、生年月日、医療記録などの患者識別情報は受け取らない 【同意取得】 CPRDに送信するデータには、個人を特定する情報が含まれていないため、CPRDとデータ共有するための患者同意は不要である（オプトアウトはあり）	【関連法規】 公衆衛生規定 【データ加工】 個人を容易に識別できるすべての情報（NIR:社会保障番号、姓名、住所等）を仮名化（英数字コードに置き換える）する必要がある

7

資料②

医療情報DBの比較 ⑤

項目	MID-NET (JPN)	Sentinel Initiative (USA)	CPRD (UK)	SNDS (FRA)
特徴 （ヒアリング等に基づく各DBの利点と欠点）	今回の調査ではヒアリングを実施していない	利点: ・DPとコミュニケーションをし、受領データをカスタマイズするなど、ある程度の柔軟性と裁量を行使できること ・SOCのQAプログラムが非常に徹底されており、詳細なチェック、傾向、比較を行うことができる	利点: ・科学的根拠に基づいた医療への貢献が可能な、充実したデータである ・患者の症例発見や全国的な診療ベンチマークを含む、診療レベルの処方と患者の安全性に関する質向上レポートが作成できる	利点: ・すべての病院の処置データが利用可能である
	今回の調査ではヒアリングを実施していない	欠点: ・規制上の問題もあり、個々の患者に関する情報を取得できていないこと ・ソースデータのレビューやDPとのコミュニケーション等に運用工数がかかる	欠点: ・手作業が多く、高度なスキルを持つ人材に依存しているため、正確なリアルタイムデータを収集することが難しい ・長期的なデータ分析は可能であるが、バイアスがかかっている。そのため、実際テキスト化すると現実世界ではうまく機能しない	欠点: ・医療機関が診療行為等のデータを登録するか否かは当該償還額と相関しているため、収集データには偏りが生じる。そのため、有効な分析を行うためにはデータ収集に関する背景等を理解する必要がある

8