厚生労働科学研究費
（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）
総合研究報告書

AIを用いた医療情報の医薬品安全への活用に向けた諸要件の調査研究

研究代表者　荒川憲昭
国立医薬品食品衛生研究所　医薬安全科学部　室長

近年、ファーマコビジランス（PV）活動においては、情報量の増大、データソースの多様化、症例評価の複雑化を背景に、人工知能（AI）技術を活用した業務支援への期待が高まっている。国際的にも、国際医学団体協議会 CIOMS をはじめ、規制当局、製薬企業、研究機関等により、PV 分野における AI 利活用の原則や課題に関する検討が進められている。本邦においても、PV 分野で AI を適切に利用するための諸要件を整理し、国内実装に向けた指針作成に資する検討が求められている。そこで本研究では、市販後安全対策および個別症例安全性報告（ICSR）を中心とする PV 業務において、AI 利活用の推進と医薬品安全管理の向上に資する知見を提供することを目的とした。

本研究では、まず、CIOMS や海外規制機関等の公開情報に基づき、PV 分野における AI 活用に関する国際的な議論や規制上の動向を把握するとともに、国内の製薬企業、IT 企業、医療機関等における検討状況について、アンケート・ヒアリング調査および文献調査を実施した。これらの結果に基づき、国内外における AI 利活用の検討状況やニーズの高い活用場面を整理し、日本の PV 実務に AI を導入する際の課題および実務上必要とされる検討事項を多角的に抽出した。

本調査研究の結果から、PV 領域で AI を利用する際の基本的な留意事項として、AI は人の最終判断を代替するものではなく、人の意思決定を支援するツールとして位置づける必要があることを示した。また、リスクベースアプローチおよび Human-in-the-Loop の概念を前提として、AI 導入プロセスにおけるリスク層別化、人による確認・監督、責任分界、ならびに透明性や説明可能性を担保するための最小安全構成を整理した。さらに、必要な医療情報ソースの要件に加え、AI 活用シミュレーションモデルを用いた検証を通じて、AI システムの品質保証、信頼性、検証方法、使用環境、ならびに法令・情報の取扱い上の留意事項についても考察した。加えて、医療機関、製薬企業、規制当局における PV プロセスの一連の情報の流れを整理し、想定される AI 使用場面ごとに、考慮すべき課題・留意点と対応策を分析した。

これらの成果は、日本の PV 実務に即した AI 活用のあり方を検討する上での基礎資料となり、今後の制度的・実務的な議論、ならびに企業・医療機関における実装検討に資するものと期待される。

研究分担者：
佐井君江（国立医薬品食品衛生研究所・医薬安全科学部）
岡本里香（神戸医療産業都市推進機構　医療イノベーション推進センター）
太田実紀（東京大学　医学部附属病院　臨床研究推進センター）
今任拓也（福岡大学　薬学部医薬品情報学）
瀬戸僚馬（東京医療保健大学・医療保健学部医療情報学科）
鈴木晶子（国際高等研究所）

A. 研究目的

近年、医薬品安全性監視活動（ファーマコビジランス：PV）の分野において、人工知能（AI）技術を活用する検討が国内外で活発に進められている。その背景には、医薬品安全性情報の取扱量の増大、データソースの多様化、症例評価の複雑化がある。特に新型コロナウイルス感染症の流行以降、個別症例安全性報告（Individual Case Safety Report：ICSR）の件数は大きく増加し、従来の人手を中心とした業務運用のみでは、膨大な安全性情報を迅速かつ一貫して処理することが困難となりつつある。PV 業務においては、文献・SNS 等からの有害事象情報の収集、症例情報の整理、重複判定、標準用語へのコーディング、ナラティブ作成、既知・未知判断、重篤性評価、因果関係評価、シグナル検出・解析など、多くの工程で AI による支援が期待されている。

一方で、PV 分野における AI 活用は、一般的な業務効率化を目的とした AI 利用とは性質が異なる。PV で扱う情報は、患者の健康被害や医薬品の安全対策に直結するものであり、AI の判断誤りや情報の見落としは、患者安全や規制上の判断に影響を及ぼす可能性がある。また、有害事象は希少事象を含むことが多く、学習データの偏りや不足、データ品質のばらつき、実運用環境における再現性の確保など、AI の性能を安定的に評価・維持するうえで特有の課題が存在する。さらに、AI の出力がブラックボックス化した場合、なぜその判断に至ったのか、なぜ重要な安全性情報を拾えなかったのかを後から説明することが困難となる。したがって、PV 領域における AI 活用では、単に処理の自動化や効率化を目指すだけでなく、人による確認・監督、説明可能性、透明性、監査可能性、責任分界を確保した実装が求められる。

国際的には、国際医学団体協議会（Council for International Organizations of Medical Sciences：CIOMS）において、2022 年 5 月に PV 用 AI に関する Working Group XIV が設置され、規制当局、産業界、研究機関、世界保健機関（WHO）等の PV および AI の専門家により、PV 分野における AI 活用の一般原則やガイダンスに関する検討が進められてきた。また、欧州医薬品庁（EMA）や米国食品医薬品局（FDA）においても、医薬品ライフサイクル全体における AI 利用、リスクベースアプローチ、人による監督、モデルの妥当性確認、データ品質、透明性、説明責任等に関する議論が進展している。本邦においても、AI 関連法令や一般的な AI ガイドラインの整備は進みつつあるが、PV 分野に特化して、AI をどの業務に、どの範囲で、どのような条件のもとで利用すべきかについては、実務に即した整理が十分とはいえない。

PV 分野における AI の利活用場面は多岐にわたり、製薬企業、医療機関、規制当局、IT 企業、アカデミアなど、関係するステークホルダーごとに期待される用途や課題に共通する部分と異なる部分があると考えられる。例えば、製薬企業では ICSR 処理や症例評価の効率化・品質向上、医療機関では副作用情報の抽出や報告書作成支援、IT 企業では AI システムの開発・提供、規制当局では大量の安全性情報の整理・評価支援などが想定される。一方で、個人情報保護、データの標準化、AI 出力の信頼性、バリデーション、責任の所在、人材・体制整備など、国内実装に向けて検討すべき課題も多い。

そこで本研究では、PV 領域における AI 活用を適切に推進するため、国内外の規制動向、文献情報、製薬企業・IT 企業・医療機関等におけるニーズや検討状況を幅広く調査し、各ステークホルダーにおける活用想定場面および課題の相違点を整理した。さらに、個別症例評価に関連する AI 活用シミュレーションを実施し、PV 業務に AI を導入する際の検証方法、品質保証、信頼性確保、人による確認の必要性等について検討した。これらの結果に基づき、国内における PV 用 AI の利活用・推進に向けた課題および留意点を整理し、今後のガイダンス案作成や実務的な導入検討に資する基礎資料を提供することを目的とした。

B. 研究方法

B-1. 国内外の医薬品安全対策における AI 活用の検討状況の調査

PV 領域における AI 活用の国内外の検討状況を把握するため、CIOMS、EMA、FDA 等の国際団体・海外規制機関が公表している文書、日本における AI 関連法令・ガイドライン、規制当局における AI 活用動向に関する公開情報を収集し、内容を整理した。

また、国内の製薬企業、IT 企業および医療機関等を対象としたアンケート調査、ヒアリング調査および文献調査により、AI 活用のニーズ、想定される利用場面、導入上の課題を整理した。

B-2. AI 活用シミュレーション

PV 領域における AI 活用の実装可能性および検証上の課題を把握するため、個別症例評価に関連する業務を想定し、「副作用の既知・未知判断支援モデル」および「段階的症例報告に基づく因果関係評価モデル」を用いた AI 活用シミュレーションを実施した。

副作用の既知・未知判断支援モデルでは、医薬品副作用データベース（Japanese Adverse Drug Event Report database : JADER）および各薬剤の添付文書情報を用いた。JADER については、解析には、令和 6 年度に研究班で作成した部分一致検索処理によるモデル v1.0 を基盤とし、令和 7 年度には GPT を用いた複合語分解・抽出機能を追加したモデル v2.0 を構築し、使用した。

B-3. 国内実装に向けた課題および留意事項の検討

国内外の調査結果およびシミュレーション解析で得られた知見を踏まえ、日本の PV 実務に AI を導入する際に必要となる

課題および留意事項を整理した。検討にあたっては、CIOMS、EMA、FDA 等の文書で示されている原則や実務上の留意点と、日本における個人情報保護法、GVP（Good Vigilance Practice：製造販売後安全管理の基準）実務、AI 関連ガイドライン等を照合した。

主な検討対象は、市販後 PV 業務、特に ICSR に関わる有害事象情報の収集、症例データの処理、評価、報告等のプロセスとした。ユースケースごとに、患者安全、規制遵守、業務影響、人による確認の必要性等の観点から、リスク層別化、責任分界、Human-in-the-Loop を前提とした確認・監督体制を検討した。

PV 領域における AI 活用を、患者安全、説明責任、透明性および監査可能性を担保しながら導入するための実務的枠組みとして検討した。

C. 結果および考察

C-1. 国内外の PV 領域における AI 活用の検討状況に関する調査

C-1-1. 各国規制機関における検討状況

<国際医学団体協議会（CIOMS）>

Pharmacovigilance（PV）への Artificial Intelligence（AI）活用に関する国際的な取り組みとして、国際医学団体協議会(CIOMS) Working Group XIV – Artificial Intelligence in Pharmacovigilance（AI in PV）が 2022 年に組織され、規制当局、産業界、研究機関、世界保健機関（WHO）の PV 及び AI の専門家により、PV 領域における AI 活用のガイダンスに関するレポート（案）が 2025 年 5 月に発表され、同年 12 月に最終化された [1]。本レポートは、生成 AI を含む AI の PV への責任ある導入フレームワークを提示するものとして、7 つの原則「リスクベース・アプローチ」「人間による監視」「妥当性・堅牢性」「透明性」「データプライバシー」「公平性と平等性」「ガバナンスと責任 」が掲げられた。重要なコンセプトとして、リスクベースによるアプローチをもとに、AI はあくまで「人間の能力を拡張・補完するツール」であり、「人間の代替ではない」点が強調されている。

<欧州医薬品庁（EMA）>

EMA では、PV を含む医薬品ライフサイクル全般への AI 活用に関するリフレクションペーパーの作成が進められ、2023 年 7 月のドラフトの公開の後、パブリックコメントを経て、2024 年 9 月に最終版が公表された [2]。ここでは、リスクベースのアプローチを基本とし、2024 年 5 月に成立した、生成 AI を含む包括的な AI 規制である「EU AI Act」[3] に基づき、患者への安全性“High Patient risk”や規制判断への影響“High regulatory impact”に応じた区分が導入された。また、欧州医薬品規制首脳会議（HMA）- EMA Big Data 運営グループ（BDSG）の「AI ワークプラン（2023-2028）」[4)]に沿って、各種ガイダンスの作成が進行し、欧州規制当局スタッフ向けの大規模言語モデル（LLMs）の利用に関するガイダンス（2024 年 8 月）[5] が作成された。その後、新たな戦略的諮問グループとして編成された HMA-EMA 合同ネットワークデータ運営グループ（NDSG）によるワークプラン「2025 年-2028 年 医薬品規制におけ

るデータと AI」[6] のもと、各種ガイダンス作成の継続とともに、AI Observatory の創設により、欧州医薬品規制ネットワークにおける AI 技術の動向の共有が推進されている[7]。また、規制当局内では医薬品開発への AI 利用ルールに関する議論や、ICSR プロセスの効率化に向けた AI ツールの開発が進んでおり、定期的な AI ワークショップの開催や、Portfolio and Technology Meeting（PTM）を通じて、企業との意見交換等も行われている[8]。

<米国食品医薬品局（FDA）>

米国 FDA では、医薬品評価研究センター（CDER）、生物製品評価研究センター（CBER）、医療機器・放射線保健センター（CDRH）との共同にて「医薬品・生物学的製剤開発における AI/ML 活用」に関するディスカッションペーパーが 2023 年 5 月に発表され、パブリックコメントを経て、2025 年 2 月に改訂版として公開された[9]。ディスカッションのテーマとして、“人間主導のガバナンス・説明責任・透明性”、“データの質・信頼性・代表性”、“モデルの開発・性能・監視・検証”が取り上げられており、本ペーパーに寄せられたコメントには、FDA の規制・監督範囲、リスクベースアプローチの運用方法、透明性と必要とされる詳細さや文書化のレベルの明確化や、国際的な調和及び医療機器（ガイダンス）との整合性、医薬品開発の関わる（機械可読）データセットの作成・共有のためのパートナーシップの確立等が挙げられていた[10]。

2023 年 10 月には、AI の安心・安全、信頼できる開発と利用に関する大統領令[11] が発令され、その中には保健福祉省に対する、医薬品開発の各フェーズを通した AI 規制戦略の確立に関する（8 つの）要求事項が含まれていた。2025 年 1 月には、医薬品規制（品質、有効性、安全性）の意思決定における AI ツールの信頼性評価に関するガイダンス「医薬品及び生物製剤の規制意思決定を支援する AI 利用の考慮事項」のドラフトが発表された[12]。

米国 FDA における PV 領域での AI 活用の取り組みとして、個別症例安全性報告（ICSR）の評価プロセス効率化のための支援ツール（InfoViP）[13]、リアルワールドデータ（レセプトや電子医療記録）を用いた統計手法の開発等も進展するほか[14]、製薬業界との情報共有の場として、EDSTM program[15] が展開している。さらに、2026 年 1 月、米国 FDA は EMA と合同にて、医薬品開発への AI の適正利用に関する 10 の原則を含む指針を発表した[16]。

<日本の AI 規制に関する動向>

日本では、AI 技術が社会や経済の発展に寄与することを目的として、「人工知能関連技術の研究開発及び活用の推進に関する法律」（AI 法）が、2025 年 6 月 4 日に公布され、同年 9 月 1 日に全面施行された[17]。日本の法規制では、EU のような厳格な規制とは異なり、「イノベーションの促進」と「リスクへの対応」の両立を基本とし、従来は事業者の自主性を重んじるソフトローが中心であったが、近年の生成 AI の台頭などの技術の進化や、国際的な議論に呼応する形で、AI 事業者ガイドラインの整備・改定が進められ、2026 年 3 月に第 1.2 版が公表された[18]。この改定で

は、「AI エージェント」や「フィジカル AI」の定義、リスクと便益が追加され、人間の判断を介在させる仕組みの設計・構築が求められている。

また、欧米規制当局内にて生成 AI 利用の進む中、日本においても、独立行政法人医薬品医療機器総合機構（PMDA）にて、生成 AI の導入が開始された[19]。なお、日本の PV 業務への利用はこれからの検討課題であり、今後、「日本における PV 分野への AI 利用に関するガイドライン」策定においては、生成 AI さらには AI エージェント等の AI 技術の急速な進展を見据えた議論が必要である。CIOMS の AI in PV レポートにおける課題とともに、国際的な AI ガイドラインとの整合性を踏まえ、製薬、臨床、規制、IT 企業を含め、それぞれのステークホルダーのニーズや課題について意見交換を重ねていくことが重要だと考えられる。

C-1-2.　アカデミアにおける検討状況

PubMed および医中誌 Web を用いて、PV 領域における AI 関連文献を調査した。その結果、日本では 2007 年に出版された文献が最も古く、その後、文献数は増加傾向にあったが、多くは学会発表などの会議録であった。一方、海外文献では 2009 年に出版されたものが最も古く、日本と同様に年々増加傾向が認められた。研究実施国としては、米国が最も多く、次いで中国が多かった。研究内容の変化を見ると、2007 年頃は副作用データのパターンを機械学習で可視化・分類する研究が中心であったが、2018 年頃からは、副作用データの解析、予測、意思決定支援など、より高度な活用へと研究の関心が移っていることが確認された。

さらに、PubMed から抽出した 13 件の科学論文を対象に、PV 領域における AI 活用に関する非定量的システマティックレビューを行った。その結果、研究内容は大きく、①副作用の予測または検出を目的とした研究、②情報の収集、抽出、補完を目的とした研究に分類された。前者では、自然言語処理を用いて臨床記録から有害事象を自動検出する研究がみられた。後者では、大規模言語モデル（LLM）を用いて、医薬品情報の抽出や要約を行う研究が確認された。このことから、AI は単にデータを処理するだけでなく、非構造化情報から安全性評価に必要な情報を取り出し、整理する手段として検討されていることが示された。

次に、ICSR に関連する業務に着目し、注目度が高いと思われるユースケース例を整理した。2026 年 1 月現在、125 件の科学論文が抽出され、LLM を用いて安全性業務との関連に基づき分類したところ、多くは「データ探索」や「シグナル探索」に関連していた。一方で、規制対応や妥当性評価に関する論文は少なく、AI を実際に PV 業務へ導入する際に必要となる運用上・規制上の検討は、まだ十分に蓄積されていないことが示唆された。

ユースケースとしては、ICSR 作成の自動化に関連して、自由記述テキストから安全性業務に重要な単語や概念を抽出すること、重複報告を自動で検知・排除すること、標準用語集への自動コーディングを行うことなどが考えられた。また、症例特性や重篤性評価に関連して、患者背景や属性

の抽出・分析、重篤性評価の自動分類なども想定された。これらは、いずれもPV業務の効率化や品質の均一化に資する可能性がある一方、最終判断をAIに委ねるのではなく、人による確認と判断を前提とした活用が必要である。

以上より、PV領域におけるAI活用に関する研究は、アカデミアにおいて進められつつあるものの、日本からの研究発信は海外と比べて少なく、世界のAI先進国に比べると検討は遅れている可能性がある。また、先行研究はデータ解析やシグナル探索に関するものが多く、規制対応、妥当性確認、責任分界、監査証跡、実運用時のリスク管理に関する検討は決して十分とは言えない。このため、今後は技術開発だけでなく、AIをPV業務のどこで、どの範囲まで使うのかを想定したうえで、ユースケースごとのリスク層別化と人による確認手順を検討する必要があることが示唆された。

C-1-3. 国内企業に関する動向調査

製薬企業及びIT企業を対象にアンケート調査により、国内のPV用AIの活用ニーズ及び技術開発・活用動向を調査した。製薬企業については、日本製薬工業協会 医薬品評価委員会 ファーマコビジランス部会に所属する75社を対象に、2024年2月21日から3月8日にかけてWebアンケートを実施し、49社から回答を得た。また、IT企業については、ライフインテリジェンスコンソーシアム（LINC）参加企業及び日本医療情報学会春季学術大会の展示企業等、計39社を対象に調査を実施し、7社から回答を得た。

製薬企業への調査では、回答企業の93.8%（46社）がPV領域におけるAI活用を推進したいと回答しており、極めて高い関心が示された。特にAIを活用したい業務としては、個別症例評価が最も多く、症例情報取得や症例評価など、人による判断が不可欠な業務においてもAI活用への期待が大きいことが明らかとなった。一方で、リスク最小化活動に対するニーズは相対的に低く、AI活用はまず、症例レベルの業務、特に個別症例評価や症例情報取得といった日常業務から進展する可能性が高いと考えられた。

想定される情報ソースとしては、個別症例評価において、副作用報告書、症例報告データベース、論文等が中心に挙げられた。特に、副作用報告書の手書き情報、テキストデータ、構造化データ、論文など、多様な形式の情報が対象となっており、非構造化データをどのように活用するかが重要な課題であることが示された。また、症例評価では、症例報告データベースに加え、論文、添付文書、医薬品リスク管理計画（RMP）などの安全性情報も重要な参照情報として位置づけられた。近年のLLMの進展を踏まえると、これまで扱いが難しかった自由記載や手書き情報の整理・抽出についても、今後は実用的な活用が期待される。

AI活用により期待される効果としては、製薬企業では、品質面では作業結果の一貫性向上、ケアレスミスの削減、評価精度の向上、報告漏れの防止が重視されていた。また、コスト面では作業工数、確認工数、外注コストの削減への期待が高く、AI活用は品質向上と業務効率化の両面から期

待されていることが分かった。一方で、規制当局への報告迅速化に対する期待は低く、これは既に定められた期限内で報告業務が運用されており、さらなる納期短縮よりも、品質と工数削減が重視されているためと考えられる。

AI 活用の現状については、2024 年の回答時点で「活用中」が 12.2%、「活用検討中」が 38.8%であり、半数以上の企業が導入または検討段階にあった。したがって、国内の PV 領域における AI 活用は、導入初期から拡大期への過渡的段階にあると考えられる。一方で、課題としては、社内スキルや専門知識の不足、バリデーション対応の不明確さ、人的資源不足、AI 精度の担保などが多く挙げられた。さらに、導入促進には、バリデーション規定や規制上の留意点を明確化したガイドライン、具体的な活用事例、人材育成、システム整備、導入・開発コストへの対応が必要とされた。

IT 企業への調査では、PV 領域における AI 関連サービスについて、「提供あり／提供予定あり」は 3 社にとどまり、アンケート時点ではサービス提供企業は限定的であった。一方で、AI 活用ソリューションの推進意向は 7 社中 5 社で認められ、関心自体は高いことが示された。IT 企業では、個別症例評価への関心に加え、リスク最小化活動やナラティブ作成など、比較的システム化しやすい領域への関心もみられた。情報ソースとしては、副作用報告書のテキストデータ、構造化データ、論文等が重視され、製薬企業と同様に非構造化データの活用が共通課題であることが確認された。

IT 企業側の課題としては、倫理的懸念やコンプライアンス対応、高い導入コスト、AI 精度の担保などが挙げられた。また、ヒアリング結果からは、現時点ではロボティック・プロセス・オートメーション（RPA）を活用した業務効率化が主流である一方、開発中のソリューションでは LLM を含む生成 AI 技術の活用が進められていることが確認された。

以上より、製薬企業と IT 企業の間には、AI 活用に対する期待や課題に一定の違いがあることが明らかとなった。製薬企業は、症例情報取得や症例評価など、日常の PV 業務に直結する領域での品質向上と工数削減を重視している。一方、IT 企業は、リスク最小化活動、ナラティブ作成、報告迅速化など、システム化・サービス化しやすい領域にも関心を示している。本研究での調査結果としては、実務面での AI 導入はまだ限定的であったものの、RPA などの自動化技術はすでに導入検討が進められており、今後、AI 技術が組み合わされることで、業務効率化がさらに進む可能性が考えられた。特に、自由記述の処理、情報抽出、重複検出、コーディング支援、要約作成などは、LLM を含む生成 AI の発展により、実用化が進みやすい領域と考えられる。

一方、製薬企業では社内スキル不足やバリデーション体制といった内部運用上の課題が大きく、IT 企業では規制対応、コスト、精度担保といった外部提供側の課題が重視されていた。今後、PV 領域で AI 活用を進めるためには、利用側である製薬企業と、提供側である IT 企業が、それぞれのニーズと課題を共有し、相互補完的に連携することが重要である。特に、個別症例評

価のようにニーズが高く、かつ人の判断が不可欠な業務では、AI の役割を明確にした上で、リスク層別化、バリデーション、責任分界、監査証跡の確保を一体として整理する必要がある。今後は、制度、技術、人材の各側面を統合的に整備し、実装可能なユースケースから段階的に導入を進めることが、国内 PV 領域における AI 普及の鍵となることが考えられた。

C-1-4.　国内医療現場に関する動向

国内医療現場における PV 領域の AI 活用可能性について、医薬品安全性情報の生成、共有、報告の実態を踏まえた調査を行った。令和 5 年度調査では、医療現場では電子カルテ情報を用いた副作用の早期検出、製薬企業では副作用症例の因果関係評価と報告、規制当局では副作用報告の情報処理や因果関係判定作業の迅速化が期待されていることが整理された。一方、医療機関では、データ品質管理や個人情報保護の課題を踏まえ、まずは RPA 等による業務支援が現実的な導入段階と考えられた。

令和 6 年度調査では、看護師および診療情報管理士を対象とした半構造化面接により、医療現場における端緒情報の把握・収集の実態を検討した。看護師からは、電子カルテや看護記録をもとに副作用や禁忌情報を自動検出し、リアルタイムで注意喚起する仕組みへの期待が示された一方、注意喚起が多すぎることにより確認作業が増えること、重要な情報が埋もれることが課題として挙げられた。診療情報管理士からは、アレルギー情報が記録されていても検出されにくいこと、データの書式が統一されていないことなどが指摘された。これらの結果を踏まえ、看護記録が医薬品安全性情報の端緒となり得るかを確認するため、匿名加工情報である看護記録データを対象とした探索的検討を行い、その結果を MEDINFO 2025 で報告した[1]。同検討では、看護記録から副作用管理に関連し得る語句を抽出できる可能性が示された一方、副作用の有無や患者状態について「Unknown」と分類される記録が多く、記録の不完全性や標準化不足により、そのまま副作用情報として用いるには課題が多いことが明らかになった。[20)]

そこで令和 7 年度は、病院薬剤部門を主たる対象として、副作用情報管理に関するアンケート調査を実施した。集計対象は 124 件であり、回答施設の多くは 400 床以上、回答者の大部分は薬剤師であった。このため、本調査結果は、一定規模以上の医療機関において副作用情報管理や報告実務を担う薬剤師の視点を中心に示すものとして解釈する必要がある。

（副作用情報管理と報告書作成の実態）

令和 7 年度のアンケート調査結果では、副作用情報の収集・管理、PMDA への報告、製薬企業への情報提供のいずれにおいても、薬剤部門が中心的役割を担っていることが確認された。したがって、医療現場における AI 活用を検討する際には、薬剤部門が担う副作用情報の収集、整理、報告書作成、確認の工程を主要なユースケースとして位置づける必要がある。

医薬品安全情報報告書の作成時には、医師記録、検査結果、投薬状況、薬剤師記録、看護記録など、複数の診療情報源が横断的に参照されていた。すなわち、副作用報告

は単一の記録だけで完結するものではなく、複数の情報を照合しながら作成されている。このため、AI 活用を検討する際には、単一文書の要約や単純な様式入力に限定せず、複数の診療情報源から必要情報を抽出し、報告様式に対応づける機能が重要である。また、AI が抽出した情報について、どの記録に基づくものかを確認できる仕組み、すなわち根拠の紐付けが必要である。

報告書作成に要する時間については、1 枚あたり 1 時間以上を要する回答が 7 割を超えており、作成者の負担としては、作成時間を確保できないこと、作成者だけでは判断できない項目が多いこと、カルテから転記する情報が多いこと、必要な情報がカルテに載っていないことが多く挙げられた。一方、報告様式そのものの使いにくさは相対的に少なかった。したがって、医療現場の負担は、入力画面の問題というよりも、情報探索、転記、要約、判断材料の整理に集中していると考えられる。AI 活用の初期的な対象としては、必要情報の抽出、経過記録の要約、報告書ドラフト作成支援などが現実的である。

AI 活用意向については、処方歴や検査結果の転記、経過記録の要約、患者基本情報の転記、副作用名の抽出に対して高い肯定率が示された。一方、被疑薬の特定についても肯定的回答は多いものの、他の項目と比べるとやや慎重な回答がみられた。このことから、医療現場では、AI に期待する作業を、情報の転記・抽出・要約と、薬学的・臨床的判断を伴う工程とに区別して受け止めていると考えられる。したがって、患者基本情報、処方歴、検査結果、経過記録などの整理・転記、報告書ドラフト作成支援を AI 活用の主な対象とし、被疑薬の特定や因果関係評価、最終的な報告内容の確定については、専門職が確認し判断する工程として残すことが適切である。

AI 活用に関する不安としては、個人情報・セキュリティ・情報管理が最も多く、次いで、正確性・誤記・信頼性、確認負担・責任所在・運用人材、事実と異なる出力・誤要約・解釈の偏り、電子カルテ連携・データ移行・システム実装が挙げられた。これらの結果から、医療現場では、AI の利便性だけでなく、出力内容を誰が確認するのか、誤りがあった場合に誰が責任を負うのか、個人情報をどのように守るのかが強く意識されていることが明らかとなった。

以上の結果より、国内医療現場における PV 領域の AI 活用は、AI が副作用報告や因果関係判断を単独で代替するものではなく、薬剤師等の専門職による確認と判断を「支援する仕組み」として位置づけることが重要である。初期導入では、薬剤部門を中心とした副作用情報管理の実態を踏まえ、情報の抽出・転記・要約、報告書ドラフト作成支援から段階的に進めることが現実的である。その際には、個人情報保護、出力内容の正確性確認、電子カルテ連携、責任分界、記録様式の標準化をあわせて検討し、AI の作業範囲、確認手順、記録方法、誤出力時の修正手順をあらかじめ定める必要があることが示唆された。

C-2. AI 活用シミュレーション

本研究では、PV 領域における個別症例評価への生成 AI 活用可能性と品質保証や

検証における課題や留意点を検討するため、令和 7 年度に、「副作用の既知・未知判断支援モデル」ならびに「段階的症例報告に基づく因果関係評価モデル」を用いた個別症例評価に関する AI 活用シミュレーションを実施した。

C-2-1. 個別症例評価 AI 活用シミュレーションモデルを用いた検討

既知・未知判断支援モデルでは、医薬品副作用データベース（Japanese Adverse Drug Event Report database：JADER）及び添付文書を用い、令和 6 年度に作成した部分一致検索処理によるモデル v1.0 を基盤として、生成 AI による複合語分解・抽出機能を追加したモデル v2.0 を構築した。解析対象は、JADER の医薬品副作用報告データセット（令和 6 年 4 月～12 月）に含まれる 8 薬剤、すなわちニボルマブ、ペムブロリズマブ、イピリムマブ、リバビリン、アピキサバン、ラモトリギン、リツキシマブ及びバンコマイシン塩酸塩とした。検討対象となった有害事象数は、薬剤ごとに 717～1,982 件であり、8 薬剤合計で 10,818 件であった。

評価にあたっては、人手で抽出した複合語を正解群とし、AI による抽出結果を True Positive（真陽性：TP）、False Positive（偽陽性：FP）、False Negative（偽陰性：FN）に分類した上で、再現率、適合率及び F 値により性能を確認した。その結果、従来の部分一致検索処理のみでは、表記ゆれ、複合語構造、上位語・下位語・同義語等を含む意味的包含関係を十分に扱えず、本来検出すべき関連事象の一部が検出されないことが確認された。これに対し、GPT を用いた複合語分解・抽出機能を追加することで、複合語に起因する検出漏れを一定程度補完できる可能性が示された。

一方で、適合率には薬剤間でばらつきがあり、一部の薬剤では誤検出の増加が認められた。F 値はニボルマブで 0.83、アピキサバンで 0.82 と比較的高かった一方、イピリムマブでは 0.29、リツキシマブでは 0.32 と低く、薬剤によってモデル性能に差がみられた。また、False に分類された事例の内訳として、FN は 11 件、FP は 50 件であり、FP はリバビリン 15 件、リツキシマブ 11 件、ペムブロリズマブ 10 件で比較的多かった。これらの結果から、添付文書における記載様式や複合語構造の違いが、AI の性能に影響する可能性が示唆された。

C-2-2. AI の品質保証・信頼性確保に関する課題

本研究では、生成 AI の追加により検出漏れを一定程度補完できる可能性が示された一方で、薬剤間で適合率や F 値に差が認められ、誤検出及び検出漏れも確認された。したがって、PV 領域で AI を用いる際には、単一の正答率や一致率のみではなく、使用目的に応じて、再現率、適合率、F 値、FN 及び FP を区別して評価する必要がある。特に FN は、安全性シグナルや重要な副作用情報の見落としにつながる可能性があるため、重点的に検証すべき項目である。一方、FP についても、業務負荷の増加要因となるだけでなく、人手による抽出漏れを補完し得る場合があるため、真の誤検出と確認結果の見直しを要する事例を区別して評価することが重要である。

また、AI の品質保証においては、性能

指標の数値確認に加え、エラー事例の分類と要因分析を行う必要がある。特に、プロンプト設計、辞書の使用方法、添付文書の記載様式、前処理・後処理ロジックが出力に与える影響を確認し、必要に応じてプロンプトチューニング、辞書ベースのフィルタリング、構文情報を用いた再判定処理を組み合わせることが求められる。さらに、判定結果だけでなく、判断根拠や参照情報を記録することで、出力の説明可能性及び信頼性を確保することが重要だと考えられた。

C-2-3. 因果関係評価における検証上の留意点

段階的症例報告に基づく因果関係評価では、症例情報の追加に伴い AI の評価が変化し得ることが示された。一方で、使用する生成 AI の応答モードにより、根拠説明の詳細さ、評価の保守性、代替原因の扱い、点数計算の整合性に差異が認められた。特に、総評価点と項目別点数の合計に不一致が生じる場合があり、AI の出力結果をそのまま採用するのではなく、項目別評価、合計点、総合評価及び判断根拠の整合性を確認する必要がある。

以上より、生成 AI は、既知・未知判断や因果関係評価における候補抽出、情報整理、判断根拠の可視化に有用である可能性があるが、最終判断を AI 単独で行うことは適切ではない。実運用にあたっては、AI を人による専門的判断を支援するツールとして位置付け、モデル、プロンプト、入力データ、辞書、添付文書の版数、前処理・後処理ロジックを記録し、変更管理及び再検証の対象とする必要がある。また、生成 AI 特有の非決定性、プロンプト依存性、モデル更新による出力変動を踏まえ、再現性確認、回帰テスト、運用後モニタリングを含む検証体系を構築することが重要であると考えられる。

D. 国内実装に向けた課題と留意事項

本研究では、上述した調査研究結果および研究班メンバーの専門性に基づき、PV 領域における AI 利活用における規制的な課題および留意事項を抽出・検討した。

D-1. 海外文書との比較と日本の課題

CIOMS、EMA、FDA 等の PV 領域における AI 活用に関する文書と国内の関連法令・ガイドライン等を照合し、日本において今後整理すべき課題を検討した。

CIOMS Working Group XIV は、PV における AI 活用について、技術仕様の詳細ではなく、長期的に活用可能な原則と良い実務（best practices）の枠組みを示している。主な柱は、リスクベース、ヒトの監督、妥当性・頑健性、透明性、データプライバシー、公平性、ガバナンスと説明責任である。

EMA は、医薬品ライフサイクル全体における AI/ML 利用を扱うリフレクションペーパーを 2024 年に公表し、開発、製造、使用、市販後を含む広い範囲で、システムライフサイクル全体を通じたリスクベースの考え方を推奨している。特に、「患者安全に影響する高リスク」と「規制判断への影響が大きい高インパクト」を区別し、申請者／製造販売業者（MAH）および開発者が規制影響とリスク分析を行うことを求めている。

米国では、AI/機械学習（ML）活用に関するディスカッションペーパーを通じて、医薬品・生物学的製剤のライフサイクル全体での AI/ML 利用を整理している。市販後安全性監視では、ICSR 処理、優先順位付け、因果関係の見込み分類、重篤性判断、集計報告の自動化などが活用可能な領域として挙げられている。

一方、日本では、2019 年の「人間中心の AI 社会原則」を起点に、AI 事業者ガイドライン、AI 法、電子カルテ標準化、医療機器・診断用 AI、個人情報処理に関するガイドライン整備が進んでいる。しかし、PV 領域に特化した AI 利活用に関する検討は十分とはいえず、実務に即した整理が今後の課題であると考えられる。

海外文書に共通する点は、AI 活用を「人間中心のリスクベース」で扱い、透明性、信頼性、説明責任を重視していることである。EMA ではシグナル検出、AE 報告の分類、重篤度スコアリングが、FDA では症例処理、症例評価、症例報告の自動化が主な PV 適用例として示されている。CIOMS もこれらを PV 向けに再整理し、リスクベースの考え方を、人の監督、性能評価、透明性、プライバシー、公平性、ガバナンスにまたがる基本原則として位置づけている。

一方で、各文書には特徴もある。CIOMS は、原則の提示に加え、運用面で「何を記録し、どう管理するか」まで踏み込んでいる。具体的には、リスク評価、テスト計画、KPI、受入基準、検証結果、緩和策、人の監督、QC、継続監視、定期再評価、事業継続計画などを文書化することを重視している。また、AI システム自体は責任主体になれないことを前提に、PV システムのオーナーが AI コンポーネントを含む品質プロセスに責任を持つこと、役割と責任を明確化すること、ベンダー監査や継続的性能監視を行うことを求めている。

EMA は、医薬品ライフサイクル全体の中で、AI/ML を規制影響および患者リスクの観点から管理する点に特徴がある。申請者／MAH が AI/ML の性能を検証・監視・文書化し、PV システムに組み込む責任を持つことを明確にしている。FDA は、現時点ではディスカッションペーパーとして、特に ICSR 処理や評価支援の自動化など、今後のユースケースと枠組み整備に向けた議論を促す性格が強い。以上を踏まえると、日本では一般的な AI 原則や関連制度は整いつつあるものの、PV 実務に即して、特に以下の点を具体化する必要があると考えられる。

第一に、PV で AI をどこまで規制・監督の対象とするか、

第二に、リスクベースの考え方をどのように運用手順へ落とし込むか、

第三に、誰が最終責任を持ち、何を記録し、どのように監査するか

第四に、他ガイドラインとどのように整合させるか、である。

EU・米国でも PV における AI 規制範囲やリスクベース運用については議論が続いており、日本では特に、実装レベルでの指針整備が重要な課題である。

D-2.　PV 領域への AI 導入の留意事項

上記の海外文書との比較から、日本では PV 領域における AI 活用について、一般的な AI 原則や関連制度を踏まえつつも、PV 実務に即した具体的な運用上の論点を整理する必要があると考えられた。そこで本研究班では、国内の PV 領域における AI 活用に関する留意事項として、現行の個人情報保護法および GVP 実務の枠内で、説明可能な形で運用するために必要となる論点を整理した。

適用範囲としては、アンケート調査およびヒアリング調査の結果を踏まえ、国内で特にニーズが高いと考えられた市販後 PV 業務のうち、主に ICSR に関わる業務および研究利用を対象とした。具体的には、有害事象情報の収集、症例データの処理、評価、報告等の各プロセスを主な対象範囲とした。

また、主に、ICSR を中心とする PV 業務への AI 活用を検討している製薬企業、医療機関、規制当局、PV 向け AI システムの開発・提供に関わる IT 企業、ならびに共同開発・研究に関わる学術研究者等が参照することを想定した。

以上の立ち位置に基づき、本研究班では、PV 領域で AI を安全かつ実務的に導入するための基本的論点として、リスクベースアプローチ、Human-in-the-Loop を前提とした人による確認・監督、ユースケースごとのリスク層別化と責任分界、ならびに最小安全構成を整理した。加えて、PV 用 AI に必要な医療情報ソースの要件、AI の品質保証・信頼性・検証法、法令、その他情報の取り扱いの留意事項を考察した。これらは、AI を単なる効率化ツールとして導入するのではなく、患者安全、説明責任、透明性および監査可能性を確保しながら活用するための基盤的な考え方である。また、これらは当局ガイダンス作成を先取りして要件を定めるものではなく、今後の制度的・実務的な議論に資する検討材料を提供することを目的に整理したものである。

D-2-1. リスク層別化と責任分界

PV 領域で AI 導入が期待される背景には、ICSR の増加と規制要件の高度化・複雑化があり、安全性情報の取扱量は一段と増大している。従来の人手中心運用だけでは高度な処理能力が必要な場面が増え、文献・SNS スクリーニング、コーディング、重複判定、ナラティブ構造化、トリアージ、シグナル検出・解析、因果関係評価支援、要約や Q&A 生成など、AI の適用範囲は拡大している。これは単なる効率化に留まらず、増大する安全性情報を適切に処理し、リスクの見逃しを防ぐための基盤整備として求められている。

図 1 は、本研究班が作成した、医療現場、企業および行政における PV プロセスマップである。PV 業務を「有害事象の発生」「副作用疑い症例の報告」「行政による収集・調査・分析」「行政による措置」の 4 つの大きなプロセスとして捉え、病院、薬局、患者、製薬企業、PMDA 等の各主体間で、どのような情報が受け渡されるのかを示したものである。各プロセスで扱われる情報の性質や利用目的は異なるため、AI に期待される用途や導入時の課題や留意点もプロセスごとに異なると考えられる。したがって、AI 活用を検討する際には、そ

れぞれのプロセスに応じた適切な技術、運用方法および留意事項を整理する必要がある。

一方で PV 用 AI は、一般業務 AI と異なり、安全性評価の誤りが患者安全に直結する。また、PV で扱う有害事象は希少となりやすく、学習データが偏りやすいことから、実運用での再現性確保は容易ではない。さらに AI がブラックボックス化すると、「なぜその判断に至ったか／なぜ重要なシグナルを拾えなかったか」を後から説明しにくくなる。PV は“低リスクゼロ”ではなく見逃しを極力避ける文化が求められるため、説明可能性の問題は特に重要であり、社会実装では精度だけでなく、患者安全・科学的再現性・制度的責任を同時に支える必要がある。

また AI 一般の課題として、非決定的挙動、誤解を招く出力、ノイズやバイアス、説明責任の難しさがあり、AI インシデントの可視化も進む。PV では患者安全・規制遵守に影響する用途ほどリスクの見立てが重要であり、導入議論は技術面に加えて事故・ハザードの視点を含めるべきである。

以上を踏まえると、AI 活用は「一律に可／不可」ではなく、ユースケースごとのリスク層別化（risk-based approach）が有効である。もし、リスクの異なる使用場面に対して一律の統制を適用すると、低リスクの用途では過剰統制による導入停滞を招き、高リスクの用途では統制不足によるシグナルの見逃し・監査不備・制度不適合を招き得る。CIOMS [1]は、リスクの程度と必要な監督の強度は、(1) 意思決定がハイステークスか、(2) AI が単独（unchecked, stand-alone）で用いられるか／人との相互作用下で用いられるか、に依存すると整理している。また、リスクベースの考え方は、人の監督、妥当性・堅牢性、透明性、公平性、データプライバシー等の設計全体に波及し、定期的に見直し・文書化されるべきである。

表 1 は、研究班で議論した各 PV プロセスにおいて想定される AI のリスク層と統制方法の基本的な考え方をまとめたものである。低リスクのユースケース（例：整形・検索等）はベースライン統制で開始しつつ、高リスク（例：重篤性/報告要否、既知/未知、因果関係、シグナル優先度など）では追加統制（影響評価、フェイルセーフ、根拠紐付け、変更管理・継続監視等）を必須化する、といった段階設計が必要となる。

また、「責任分界（責任分解）」も同様に必要となる。AI システム自体は責任主体になり得ないため、誰が何に責任を持つかを事前に切り分け、説明・是正・再発防止が可能な状態にする必要がある。CIOMS レポートでは [1]、AI システムを配備・運用・管理して PV 義務を果たす主体に対して、役割・責任・（必要に応じて）法的責任を明確にすることを「アカウンタビリティ」として位置づけ、規制当局、ベンダー、利用者、開発者、データ提供者等を含む関係者に適切なガバナンス措置を求めている。

さらに、PV システムの品質プロセスと AI 要素の監督はシステムオーナー（運用主体）が負い、ベンダーとの協働、監査・適格性確認、継続的な性能監視等により“inspection ready”を確保すること、AI の版管理・トレーサビリティや関連文書（例：

PSMF（Pharmacovigilance System Master File）での位置づけが重要であることが示されている。

加えて国内の一般的なAIガイドラインや法制度は整備が進む一方、PV実務に必要な使用基準・監査・説明責任の具体化はなお抽象的であり、PV実装で必要な運用上のレッドラインを具体化する観点からも、責任分界の明確化が求められる。

＜責任分界の例＞

1) 提供者（ベンダー／開発側）
仕様・制約・想定利用（Context of Use）、評価結果、更新情報、セキュリティ要件、ログ取得可能範囲等を明示し、変更時の影響と必要対応を提示する。
2) 利用者（医療機関／製薬企業／規制当局）
ユースケース適合性の判断、HITLを含む運用設計、監査証跡の確保、継続監視と是正、個人情報保護・GVP遵守について最終責任を負う。CIOMSのユースケース整理でも、最終的な規制上のアカウンタビリティは、内容をレビューし検証する専門家／ユーザー側に帰属する旨が示されている。
3) 組織内（RACI等）
RACI (Responsible、Accountable、Consulted、Informed：実行責任者、説明責任者、協議先、報告先) の考え方等に基づき、組織内の役割と責任を明確化する必要がある。具体的には、プロセスオーナー（業務責任）、医学専門家（判断責任）、QA（監査・逸脱管理）、IT/セキュリティ（基盤・アクセス・ログ）、データ保護担当（DPIA：Data Protection Impact Assessment、データ保護影響評価等の個人情報・データ保護対応等）、データサイエンス（評価設計、性能評価、監視指標の設定等）の役割と責任を整理することが重要である。特に、AIの出力が安全性判断や規制対応に影響し得る高リスク用途では、業務部門から独立したレビューや承認プロセスを設けるなど、確認体制を強化することが望ましい。

PV領域でAIを安全かつ実務的に導入するには、(1) ユースケースごとのリスク層別化により、見逃しリスク・規制影響の大きい用途ほど統制を強化すること、(2)責任分界により、監査可能性・再現可能性を担保し、AI関与下でも説明責任を果たせる運用（記録保持、根拠紐付け、変更管理、継続監視）を確立することが重要である。

D-2-2. 最小安全構成

「最小安全構成」とは、PV領域におけるAI利活用の初期導入段階において、リスクを統制可能な範囲に限定するための基準モデルである。これは、AI利活用の多様な形態の中で、比較的リスクを限定し得る構成要素を抽出し、基準点として整理したものである。具体的には、① 閉域運用、② 非学習・非保持、③ 人の判断記録という三要素から構成される。これら三要素は、それぞれ「データ統制」「情報拡散防止」「責任の可視化」という異なるリスク次元に対応する。

第一に、閉域運用が挙げられる。医療

データを含む AI 利活用においては、データの外部送信や不必要なアクセス拡大が重大なリスクとなる。このため、クラウド利用の可否を含め、データの所在および流通経路を制御可能な環境下で運用することが基本となる。閉域的な運用は、単に情報漏えいリスクを低減するにとどまらず、アクセス主体の限定を通じて責任の所在を明確化する機能を有する。

第二に、非学習・非保持である。AI モデルが利用時の入力データを次の学習に用いないこと、また入力・出力データを継続的に保持しないことは、個人情報保護およびデータガバナンスの観点から重要である。特に PV 領域では、患者情報が含まれる可能性があるため、モデル内部への情報混入や再利用のリスクを回避する設計が求められる。この点は、契約上においても明確に担保される必要がある。

第三に、人の判断記録である。AI の出力はあくまで判断支援として位置づけられ、最終判断は人間が行うことが前提となるが、重要なのはその判断過程が記録されることである。すなわち、AI の出力、人間の評価、最終判断の関係が追跡可能な形で残ることにより、責任の明確化および事後検証が可能となる。この点は、監査可能性および説明責任を担保する基盤となる。

さらに、PV 領域においては、提出文書における説明可能性が要求されることから、AI 出力とその根拠情報との紐付け・トレーサビリティ（traceability）が不可欠である。これは、ブラックボックス的な判断を排し、外部的な証跡によって判断の妥当性を説明可能とする構造を意味する。

また、データ分布や利用環境の変化に伴う性能変動（いわゆるドリフト）への対応も重要である。すなわち、変更管理および継続的監視の仕組みを通じて、AI の挙動が意図しない形で変化することを防止し、その影響を適切に評価することが不可欠である。

以上の最小安全構成は、すべての AI 利用に一律に適用される規範ではないが、国内実装における初期的な基準点として位置づけることができる。とりわけ、高リスク又は高い規制影響を伴う場面においては、これらの要素に加え、より厳格な統制（例：独立した検証プロセス、利用制限、追加的監査等）を講じることが求められる。

このように、PV 用 AI における基本原則は、「人間責任」「監査可能性」「データガバナンス」といった観点を中核としつつ、技術的構成および運用設計の双方において具体化されるべきである。これらは、最終的に「人間責任」を中核とし、それを支える要素として「監査可能性」および「データガバナンス」が位置づけられる。最小安全構成は、そのための実務的な出発点を提示するものだと考えられる。

D-2-3. 医療情報ソースの要件

(1) AI に求められる医療情報の質

AI は、副作用の自動検出や因果関係評価の効率化など、従来人手を要していた PV 業務の質的・量的向上に大きく貢

献し得る。一方で、AI の判断精度は、入力される医療情報の粒度、品質、時系列的整合性に大きく左右される。すなわち、AI の性能そのものだけでなく、どのような情報を、どのような形で入力するかが結果を決定づける。「Garbage in, garbage out」の原則は、患者安全に直結する PV 活動において特に重要であり、AI の誤判断が患者安全に影響し得ることを踏まえると、情報ソースの要件を明確化し、データの信頼性と完全性を確保することが不可欠である[21]。

現在の PV 実務では、因果関係評価や安全対策の要否を判断するために必要な情報が十分に揃っていない症例も少なくない。この状態のまま AI を導入し、症例報告の処理速度だけを高めても、判断の前提となる情報不足は解消されない。むしろ、不十分な情報に基づく「もっともらしい判定結果」が大量に生成され、リスクコミュニケーションを歪める可能性がある。したがって、AI 時代の PV の信頼性を支えるためには、情報の質・量・入手経路を整理し、どの情報を、誰が、どの場面で収集し、どの形式で保持し、どのようなルールで AI の学習・推論に用いるかを標準化する必要がある。

本邦の PV プロセスは、おおむね、
①有害事象の発生（患者の病院受診から医薬品の処方と服用、有害事象発生）
②副作用疑い症例の報告（医療機関や製薬企業による情報収集、規制当局への報告）
③行政による収集・調査・分析
④行政による安全対策措置
の流れで構成される。このうち、①～②の発生時点での記録や院内報告など初期段階の情報が不足すると、③～④の分析や安全対策も、不十分な材料にて行われることになる。PV に必要な情報は、有害事象発生以前の段階（原疾患による受診、処方、服薬）からすでに収集が始まっており、日常診療における情報の質が最終的な副作用報告の質に大きく影響することを認識する必要がある。

(2) PV に必要な情報と AI 活用時に必要な情報の違い

AI を用いた PV 活動を適切に進めるためには、「PV に本来必要な情報」と「AI を活用する際に追加的に必要となる情報要件」とを区別して考える必要がある。前者は、人間による従来の評価にも必要な基本的な安全性情報であり、後者は AI による学習や推論の精度、再現性、信頼性を確保するための留意点である。

(3) PV に本来必要な情報

ICH E2D では、ICSR に必要な最低限の情報として、「特定された報告者」「特定された患者」「副作用」「被疑薬」の 4 要素を示している。また、症例経過には、患者情報、治療内容、既往歴、有害事象の臨床経過など、関連する臨床情報の要約が必要とされる[22]。

因果関係評価では、一般に、服薬開始日・有害事象発現日・中止日・回復日などの時間的関連、被疑薬の用法・用量・投与経路・投与期間・適応などの薬剤情報、薬剤中止後の転帰や再投与後の再発の有無などの dechallenge / rechallenge 情報、用量依存性、検査値・臨床所見、合併症・併用薬・既往歴などの代替原因、

添付文書や公知情報との照合、転帰情報などが重要となる。特に、陰性所見は因果関係評価において重要である。所見が認められなかったという事実も評価に必要な情報であり、省略すべきではない。記載がない場合、それが未確認なのか、欠損なのか、陰性なのかを判断できず、情報不足として扱われる可能性がある。したがって、有害事象に関連する因子については、陰性所見も明確に記録することが望ましい。

- 時間的関連：服薬開始日・有害事象発現日・中止日・回復日など
- 被疑薬の詳細：用法・用量・投与経路・投与期間・適応（使用理由）など
- Dechallenge/Rechallenge の情報：薬剤中止後の転帰、再投与後の再発の有無など
- 用量と反応の関係：用量依存性の有無など
- 検査値・臨床所見：有害事象の診断妥当性、重症度評価など
- 代替原因の有無：合併症、他の薬剤、既往歴等による説明可能性など
- 既知の薬理・安全性情報：添付文書、公知情報との照合など
- 転帰情報：回復・後遺症・死亡の別、入院要否など

(3) AI 活用時に特有の情報要件と留意点

AI を PV に用いる際には、従来の安全性評価に必要な情報に加え、AI 特有の留意点がある。

第一に、入力データの網羅性と偏りである。AI は、訓練データや入力データに含まれる偏りを学習し、場合によっては増幅する。特定の人種、地域、年齢層、社会経済層が少ないデータで学習されたモデルでは、その集団に関するシグナルを見落としたり、誤ったシグナルを検出したりするリスクがある[23]。これは従来の手動評価でも問題となり得るが、AI が大量データを自動処理する場合には、偏りが組織的・系統的に増幅される可能性があるため、特別な注意が必要である[24]。

第二に、非構造化データの扱いである。電子カルテの自由記載には、重要な臨床情報が多く含まれており[25]、NLP や機械学習により有用な情報を抽出できる可能性がある[26]。一方、自由記載は、施設、医師、システムによって記載様式が異なり、略語、誤字、省略表現も含まれる。このため、標準化されていないまま AI を適用すると、情報抽出の精度にばらつきが生じる[27,28]。さらに、日本語では主語の省略、否定表現、時系列表現、文脈依存的な記載が多く、日本語特有の表現構造にも配慮が必要である[29,30]。

第三に、欠損値への対応である。AI モデルは、欠損した情報を学習データから推定して補完する場合がある。特に LLM を含む AI システムが、不明な臨床情報を「それらしく」生成してしまうことは、副作用評価や因果関係評価において重大なリスクとなる[23]。例えば、転帰、dechallenge の有無、既往歴などが不明であるにもかかわらず、AI が推測で補完すると、評価の信頼性は大きく損なわれる。したがって、欠損値は AI が補完するのではなく、欠損値であることを明示し、必要に応じて追加情報収集を促す仕様

とすることが重要である。AI は情報を「埋める」ものではなく、情報の不足を可視化し、人による確認や追加収集を支援するものとして位置づけるべきである。

D-2-4. AI の品質保証、信頼性、検証法

本項では、まず CIOMS レポート[1)]で示されている PV 領域における AI 活用の品質保証・信頼性確保に関する基本的考え方を踏まえた上で、本研究の AI 活用シミュレーションモデルの PoC 検証で確認された具体的課題を整理し、これらを基に、バリデーション、性能評価、ドリフト監視、使用環境及びデータガバナンスに関する実装上の留意点を検討した。

CIOMS Working Group XIV による PV 領域における AI の活用に関する報告書では、リスクベースアプローチ、人間による監督、妥当性・堅牢性、透明性、説明可能性、データプライバシー、ガバナンス及び説明責任が重要な原則として整理されている。[1)] また、AI は人による意思決定を代替するものではなく、適切な監督、検証及び管理の下で意思決定を支援するものとして位置付ける必要があるとされている。

C-2 の項で述べたように、本研究では、AI 活用シミュレーションモデルの PoC 検証として、副作用の既知・未知判断支援モデル及び段階的症例報告に基づく因果関係評価モデルを検討した。既知・未知判断支援モデルでは、部分一致検索処理や辞書照合を主体とするルールベース型処理に、大規模言語モデル（large language model : LLM）による自然言語理解・複合語分解を組み合わせることで、従来法では検出困難であった複合語や意味的包含事象を補完的に抽出できる可能性が示された。一方で、薬剤間で適合率や F 値に差がみられ、False Positive（偽陽性 : FP）及び False Negative（偽陰性 : FN）も認められた。また、段階的症例報告に基づく因果関係評価では、症例情報の追加に伴い評価が変化し得る一方、使用するモデル又はモードにより、評価の保守性、代替原因の扱い、点数計算の整合性に差異が生じることが確認された。

これらの結果から、PV 領域における AI システムの品質保証においては、AI を単一の自動判定システムとして扱うのではなく、利用目的、モデル特性及び出力結果が業務判断に与える影響に応じて、バリデーション、性能評価、運用後監視及びデータガバナンスを組み合わせて検討する必要があると考えられる。

（1）バリデーション法

既知・未知判断支援モデルのうち、部分一致検索処理や辞書照合を主体とする処理では、検索対象、照合ルール、辞書、出力条件が比較的明確であるため、従来のコンピュータ化システムバリデーション（computerized system validation : CSV）に準じた検証を適用しやすいと考えられる。具体的には、仕様通りに検索・照合・候補提示が行われること、辞書更新、添付文書改訂、ルール変更が出力結果に与える影響を確認することが中心になると考えられる。

一方、LLM を用いた複合語分解、意味補完又は因果関係評価支援では、決定論的な検証のみでは十分でない可能性がある。LLM の出力は、プロンプト、入力情報の形式、モデル又はモード、モデル更新等により変動し得るため、同一入力に対する再現性、人的評価との一致性、判断根拠の妥当性、出力の安定性を含めて検証することが望ましい。したがって、LLM 主体又は LLM を組み込んだモデルでは、単純な正解一致だけでなく、評価根拠、判断過程、出力の説明可能性及び変更時の影響評価を含めた検証が重要になると考えられる。

（2）性能評価法

性能評価では、AI が担う役割と許容可能なリスクを明確にした上で、評価指標を設定することが重要である。既知・未知判断支援モデルでは、LLM の導入により検出漏れを補完できる可能性が示された一方、誤検出の増加や薬剤間の性能差も認められた。このため、単一の正答率や一致率のみで評価するのではなく、True Positive（真陽性 : TP）、FP、FN を区別し、再現率、適合率、F 値等を組み合わせて評価することが有用と考えられる。

特に、FN は本来検出すべき副作用情報の見落としにつながる可能性があるため、安全性評価上のリスクとして重点的に確認すべき項目である。一方、FP は業務負荷を増加させる可能性があるが、本研究では人手による抽出漏れを補完し得る事例も含まれていた。したがって、FP を単純な誤検出として扱うのではなく、真の誤検出と、人手確認結果の見直しを要する事例とを区別して評価することが望ましい。

因果関係評価支援では、静的な正答率だけでなく、症例情報の追加に応じて評価がどのように変化するかを確認することが重要と考えられる。具体例として、時系列の整合性、デチャレンジ、再投与、代替原因、Naranjo スコア等との整合性、同一症例に対する評価の一貫性、評価根拠の明示などは、主要な評価項目になり得る。また、総評価点と項目別点数の合計に不一致が生じる可能性があるため、AI 出力をそのまま採用するのではなく、項目別評価、合計点、総合評価及び判断根拠の整合性を確認することが望ましい。

（3）ドリフト監視

運用後には、モデル、入力データ、参照情報又は業務環境の変化に伴う性能低下、すなわちドリフトへの対応が重要となる。CIOMS レポートでも、導入済みモデルの詳細な文書化、モデル構造、訓練・テストデータ、性能評価結果の記録に加え、導入後の継続的改善やモデルドリフトを含む性能劣化の監視が重要とされている。

既知・未知判断支援モデルでは、添付文書の改訂、注意喚起記載の様式変更、辞書更新、複合語パターンの変化が出力結果に影響し得る。このため、辞書や添付文書等の参照データを更新する際には、影響評価を行い、必要に応じて同一検証データセットを用いた再評価を実施することが望ましい。

因果関係評価支援モデルでは、症例分布の変化、新規モダリティや適応拡大に伴う副作用傾向の変化、医学知識や診療ガイドラインの更新、さらに LLM 自体の更新が評価結果に影響する可能性がある。そのため、代表的な症例、境界事例、情報不足事例、代替原因を有する事例等を含むベンチマーク症例セットを設定し、定期的に同一条件で再評価することが有用と考えられる。あわせて、プロンプト、モデルバージョン、入力データ、辞書、添付文書版数、前処理・後処理ロジックを記録し、変更管理及び再検証の対象とすることが望ましい。

（4）使用環境とデータガバナンス上の留意点

クラウド環境で AI システムを利用する場合には、症例情報が外部環境へ送信されるリスクを評価する必要がある。特に、入力データがモデルの再学習に利用されないこと、入出力データが保持されないこと、保存される場合には保存期間及び利用目的が明確であること、契約条件やデータ取扱いポリシーが PV 業務に適合していることを確認する必要がある。

ローカル LLM は、外部送信を避けられるため高いデータ保護性が期待できる一方、モデル性能の維持、更新、セキュリティ対策、運用保守を自組織で担う必要があり、人的・技術的負担が課題となる可能性がある。また、品質保証及び監査対応の観点からは、入出力ログの保存が有用である。ただし、ログには症例情報や個人情報が含まれる可能性があるため、マスキング又は匿名化、保存期間、アクセス権限、削除方針を明確化する必要がある。 Retrieval-Augmented Generation（RAG）を利用する場合には、知識ベースに格納するデータの範囲、更新頻度、アクセス権限、削除方針の管理も重要である。

以上より、PV 分野における AI システムを実装する際には、AI を最終判断の主体ではなく、人による専門的判断を支援するツールとして位置付けることが基本となる。あわせて、モデル特性に応じたバリデーション、性能評価、ドリフト監視、ログ管理、非学習・非保持を含むデータガバナンスを一体として設計することが、AI 出力の信頼性、説明可能性及び監査可能性を確保する上で重要と考えられる。

D-2-5. 法令、その他情報の取り扱いの留意事項

PV 用 AI は、技術的な適用可能性のみならず、既存の法令および規制体系との整合の中で位置づけられる必要がある。特に、日本においては、個人情報保護法（APPI）および医薬品の安全性監視に関する規制（GVP）との関係を踏まえた運用設計が不可欠である。

まず、個人情報保護の観点からは、AI 利活用に伴うデータ取扱いが、利用目的の特定、データ最小化、安全管理措置といった基本原則に適合する形で構成される必要がある。特に、AI システムへの入力データについては、個人データの過剰な利用を回避する観点から、マスキングや仮名化等の措置を講じることが重

要となる。また、複数のデータソースを結合して分析を行う場合には、再識別リスクが高まることから、結合の必要性および方法について慎重な検討が求められる。

次に、GVP との関係では、PV 業務における判断の適正性および記録の保存が重要な要件となる。この点において、AI の利活用は、従来の業務プロセスに新たな層を加えるものであり、AI 出力の利用方法、最終判断の主体、判断過程の記録の在り方を明確にしておく必要がある。特に、AI を用いた分析結果を安全性評価や報告に用いる場合には、その根拠となるデータおよび処理過程が追跡可能であること、すなわち監査証跡が確保されていることが肝要であると思われる。

また、AI サービスの利用に際しては、委託先管理の重要性が一層高まる。PV 業務のすべて、または一部を CRO（開発業務受託機関）等の外部委託先に委託する場合、当該委託先が利用する AI ツールやデータ取扱いの条件が、最終的なリスクに直接影響する。このため、契約において、データの取扱い、学習への利用の有無、ログの管理、再委託の範囲等を明確化するとともに、必要に応じて監査権限を確保することが重要である。なお、業務を委託する場合であっても、最終的な責任は委託元に帰属する点に留意する必要がある。

さらに、AI 利活用に伴うログ管理は、説明責任および規制対応の観点から重要な役割を果たす。すなわち、どのデータが用いられ、どのような処理を経て、いかなる判断がなされたかを記録することにより、事後的な検証および当局対応が可能となる。一方で、ログには個人情報や機微情報が含まれる可能性があるため、その保存範囲および保存期間については、データ最小化の観点から適切に設計される必要がある。

以上のように、PV×AI の実装においては、技術的な可否判断にとどまらず、APPI および GVP といった既存制度との整合を前提とした運用設計が求められる。すなわち、AI の利用は制度を置き換えるものではなく、既存の規制枠組みの中でその適用様式を変容させるものである。その際には、データ管理、責任分担および監査可能性を中核とするガバナンスの確立が不可欠だと考えられる。

D-3. 想定ユースケースにおける留意点

以下では、本研究で実施した国内外動向調査、アンケート・ヒアリング調査および AI 活用シミュレーションの結果を踏まえ、本研究班内で抽出・議論したユースケースごとに、リスク層別化ならびに統制・責任設計上の主な留意点を整理した。

(1) 副作用報告のデータベース登録（入力案提示支援）

副作用報告のデータベース登録は、AI 導入による効率化効果が期待される一方、後続の安全性評価や規制対応の基盤データを扱うため、中リスク程度のユースケースとして扱うことが妥当と考えられる。誤抽出や誤分類が生じた場合、データ品質の

低下を通じて、集積解析やシグナル検出にも影響し得る。

このため、AI の役割は登録そのものではなく、入力案の提示や転記支援に限定し、最終登録は人が確認した上で行う設計が望ましい。対策としては、HITL 設計、QC 担当者による修正・確定、根拠となる原文箇所の提示、履歴保存、閉域運用・マスキング、属性別性能評価、偏り分析等を組み合わせることにより、説明可能性と監査可能性を確保しやすくなると考えられる。

(2) 副作用疑い症例の初期評価（Evidence Pack／論点整理）

副作用疑い症例の初期評価では、時系列整合性の確認、代替要因の洗い出し、不足情報の抽出、関連根拠の提示などを通じて、医学専門家による一次評価を支援することが想定される。入力情報としては、ICSR 構造化項目に加え、原資料、添付文書、社内照会情報、過去症例、シグナル情報など、複数の情報源を扱う可能性がある。

このユースケースでは、AI は情報の構造化・抽出、矛盾点の提示、代替要因や不足情報の整理、類似症例や根拠候補の提示などを担うことが考えられる。ただし、AI は結論を出す主体ではなく、最終的には安全管理責任者や医学専門家が該当箇所を確認し、人の判断により初期評価を行う設計が適切と考えられる。特に、根拠紐付け、Evidence Pack 化、不確実な場合に人へ回すフェイルセーフ設計を強化することにより、説明責任と見逃し抑制に資すると考えられる。

(3) 文献・外部情報からの安全性情報抽出（翻訳・候補化・優先度付け）

文献・外部情報からの安全性情報抽出では、学術論文、抄録、非構造化文書、SNS、規制当局公開情報等を対象に、副作用名、薬剤名、重篤性、症例情報などを抽出することが想定される。また、非英語文献の翻訳、MedDRA 候補化、既存データベースとの照合、類似症例候補の提示、要約や優先順位付けなどにも AI 活用の余地があると考えられる。

一方で、抽出結果は患者安全に関わる判断材料となり得るため、人による最終レビューを組み込むことが望ましい。特に、RAG 等を用いて参照元を明確化し、原文と AI 抽出結果を照合した上で、専門家が採否を判断する設計が有用と考えられる。AI による要約や翻訳についても、原文に戻って確認できる仕組みを設けることが望ましい。

(4) 既知・未知判断支援（予測性判断）

既知・未知判断支援は、報告された有害事象が添付文書等の注意喚起に記載された事象に該当するかを確認し、予測性判断を支援するユースケースである。規制当局への報告要否や安全性評価の位置づけに影響し得るため、中〜高リスク程度のユースケースとして扱うことが考えられる。

本ユースケースでは、部分一致検索や辞書照合など、比較的説明しやすいルールベース処理を主軸とし、LLM は複合語分解や意味的に近い表現の補完に限定する設計が望ましいと考えられる。AI が提示するのは「該当候補」や「関連可能性」にとどめ、既知・未知の最終判断は PV 担当者

または医学専門家が行うことが適切である。

性能評価では、候補抽出段階では見逃しを抑えるため再現率を重視し、偽陽性については人手確認で制御する二段階設計が有用と考えられる。また、参照した添付文書の版管理、辞書・シソーラスの更新管理、AI が提示した候補と根拠箇所の記録、最終判断者と判断理由の記録が求められる。添付文書改訂や辞書更新により結果が変わる可能性があるため、変更管理と再評価も運用に組み込むことが望ましい。

(5) 因果関係評価支援

因果関係評価支援は、症例情報、時系列情報、被疑薬、併用薬、既往歴、検査値、dechallenge/rechallenge、代替要因等を整理し、専門家による因果関係評価を支援するユースケースである。評価結果が安全対策や規制判断に影響し得るため、高リスクに近いユースケースとして扱うことが妥当と考えられる。

本ユースケースでは、AI が因果関係を最終判定するのではなく、評価に必要な情報の整理、不足情報の抽出、時系列の可視化、代替要因の提示、Naranjo スコア等の既存評価軸との対応づけ、根拠候補の提示を担う位置づけが適切と考えられる。特に ICSR は初期段階で情報が限られ、その後の追加情報により評価が変化し得るため、AI 出力は「現時点の情報に基づく支援結果」として扱い、情報更新に応じて再評価する設計が望ましい。

統制としては、同一入力に対する出力の安定性確認、症例情報の追加に伴う評価変化の妥当性確認、人手評価との一致度確認、根拠箇所の明示などが考えられる。また、不明情報を AI が推測で補完することは避け、欠損情報は欠損として明示し、追加情報収集につなげる設計が望ましい。最終判断は安全管理責任者や医学専門家が行い、AI 出力、人による修正、最終判断理由を記録することで、説明可能性と監査可能性を確保しやすくなると考えられる。

(6) シグナル検出・シグナル評価の優先度付け

シグナル検出やシグナル評価の優先度付けでは、症例報告データベース、文献、外部情報、過去症例、RMP、添付文書等を横断的に参照し、注目すべき組み合わせや評価対象候補を抽出することが想定される。見逃しが患者安全に影響し得る一方で、偽陽性が増えると評価負担も増大するため、高リスク寄りのユースケースとして扱うことが考えられる。

AI は、集積傾向の可視化、類似症例の候補提示、文献・外部情報との照合、優先確認すべき論点の提示などに有用である可能性がある。ただし、シグナルの有無や安全対策の必要性を AI が単独で判断する設計は避けるべきであり、PV 専門家が、AI の提示した候補を既存の科学的知見、疾患背景、曝露状況、代替要因、既知リスク等と照合し、最終評価を行うことが望ましい。

統制としては、使用データソースの範囲、抽出条件、閾値、モデル・プロンプト・辞書の版、候補提示の根拠、採否判断の記録を残すことが重要と考えられる。また、症例分布や報告傾向の変化、製品の適応拡大、新規モダリティの登場により性能が変化

し得るため、継続的なドリフト監視と定期的な再評価を行うことが望ましい。

E. まとめ

本研究では、PV 分野における AI 活用に関して、CIOMS や海外規制機関等における国際的な議論と規制動向に加えて、国内の製薬企業、IT 企業、医療機関等における AI 活用の検討状況や文献などの調査を実施した。これらの結果に基づき、国内外における AI 利活用の動向、ニーズの高い活用場面、導入上の課題を整理し、日本の PV 実務に AI を導入する際に必要となる実務上の検討事項を多角的に抽出した。

PV 領域で AI を安全かつ実務的に導入するには、医療現場で発生した情報が、PV/ICSR 業務に必要な形で適切に整理され、後続の評価に利用できる情報基盤として整備されていることが前提となる。その上で AI を活用すれば、情報抽出、要約、照合、論点整理などを通じて、業務効率化だけでなく、見落としの低減、評価の一貫性向上、判断根拠の可視化など、より質の高い PV 業務に繋がることが期待される。

一方で、PV は生命・健康に直結する分野であり、AI の誤判断や見逃しは患者安全に影響し得る。そのため、予防原則に基づく慎重な対応は重要であるが、過剰な慎重さは安全対策の迅速化や業務効率化を妨げる可能性もある。したがって、AI 導入にあたっては、科学的根拠と倫理的判断のバランスを取りつつ、ユースケースごとのリスク層別化により、見逃しリスクや規制影響の大きい用途ほど統制を強化することが重要である。

また、PV×AI に求められるのは「完璧な自動判断」ではなく、人間との相互補完と、システム全体としての堅牢性の確保であると考えられる。AI は責任主体にはなり得ず、「重篤性」「有害性」「因果関係」「安全対策の要否」などの価値判断は、最終的には人が担う必要がある。そのため、Human-in-the-Loop を前提に、AI は意思決定支援ツールとして位置づけるべきである。

さらに、生成 AI 等では、「もっともらしい」誤情報や合成情報が混入するリスクもあるため、情報の出所、根拠、症例報告の出典、確認履歴を追跡できることが重要である。責任分界を明確にし、記録保持、根拠紐付け、変更管理、継続監視を通じて、AI 関与下でも説明責任と監査可能性を担保する必要がある。

以上より、PV 領域における AI 導入は、利便性や効率性のみで正当化されるものではなく、透明性、説明責任、公共的正当性を確保した上で進める必要がある。本研究成果は、日本の PV 実務に即した AI 活用のあり方を検討するための基礎資料であるとともに、今後の制度的・実務的な議論に資するものと考えられる。また、企業・医療機関における自主的な運用ルール、社内ガイドライン、教育・啓発などが整えられていくことで、人と AI が相互補完する形で段階的に実装していくことが望ましいと思われる。

F. 参考文献

1) CIOMS : Artificial Intelligence in Pharmacovigilance https://cioms.ch/artificial-intelligence-inpv/#description

2) EMA. Reflection paper on the use of Artificial Intelligence (AI) in the medicinal product lifecycle https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-use-artificial-intelligence-ai-medicinal-product-lifecycle_en.pdf
3) The EU Artificial Intelligence Act. https://artificialintelligenceact.eu/
4) Multi-annual AI workplan 2023-2028 HMA-EMA Big Data Steering Group : https://www.ema.europa.eu/en/documents/work-programme/multi-annual-artificial-intelligence-workplan-2023-2028-hma-ema-joint-big-data-steering-group_en.pdf
5) EMA . Guiding principles on the use of large language models in regulatory science and for medicines regulatory activities https://www.ema.europa.eu/en/documents/other/guiding-principles-use-large-language-models-regulatory-science-medicines-regulatory-activities_en.pdf
6) Joint HMA/EMA Network Data Steering Group. NDSG workplan 2025-2028 Data and AI in medicines regulation. version 1.0 https://www.ema.europa.eu/en/documents/other/network-data-steering-group-workplan-2025-2028_en.pdf
7) EMA. 2024 AI Observatory report https://www.ema.europa.eu/en/documents/report/2024-ai-observatory-report_en.pdf
8) EMA. Portfolio and technology meetings https://www.ema.europa.eu/en/human-regulatory-overview/research-development/portfolio-technology-meetings
9) FDA. Using Artificial Intelligence & Machine Learning in Development of Drug and Biological Products : Discussion Paper and Request of Feedback : https://www.fda.gov/media/167973/download?attachment
10) FDA. Regulatory Education for Industry (REdI) Annual Conference（May 29th, 2024）. https://sbiaevents.com/redi2024/
11) A Presidential Document by the Executive Office of the President on 11/01/2023 https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence
12) FDA. Considerations for the Use of Artificial Intelligence To Support Regulatory Decision-Making for Drug and Biological Products Draft Guidance for Industry and Other Interested Parties https://www.fda.gov/media/184830/download
13) Spiker J, et al.. Information Visualization Platform for Postmarket Surveillance Decision Support Drug Saf. 2020;43:905-915.
14) Desai RJ, et al. Broadening the reach of the FDA Sentinel system: A roadmap for integrating electronic health record data in a causal analysis framework.NPJ Digit Med. 2021;4:170.
15) FDA. Emerging Drug Safety Technology Meeting (EDSTM) program. https://www.fda.gov/drugs/cder-emerging-drug-safety-technology-program-edstp/emerging-drug-safety-technology-meetings-edstm
16) FDA. Guiding Principles of Good AI Practice in Drug Development https://www.fda.gov/about-fda/artificial-intelligence-drug-development/guiding-principles-good-ai-practice-drug-development
17) 内閣府. 人工知能関連技術の研究開発及び活用の推進に関する法律（ＡＩ法） https://www8.cao.go.jp/cstp/ai/ai_act/ai_act.html
18) 経済産業省. AI 事業者ガイドライン（第 1.2 版） https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/20260331_report.html
19) 独立行政法人 医薬品医療機器総合機構. PMDA 業務への生成 AI の利用開始について（令和８年４月） https://www.pmda.go.jp/files/000280170.pdf
20) Wakabayashi S, Seto R. Extracting Pharmaceutical Safety Information from Nursing Records: Utilizing ChatGPT for Data Categorization. In: MEDINFO 2025 — Healthcare Smart × Medicine Deep. Studies in Health Technology and Informatics. 2025;329:352-356. doi:10.3233/SHTI250860.

21) Teno JM. Garbage in, Garbage out—Words of Caution on Big Data and Machine Learning in Medical Practice. JAMA Health Forum. 2023;4(2):e230397–e.
22) 厚生労働省医薬食品局安全対策課長通知. 承認後の安全性情報の取扱い：緊急報告のための用語の定義と報告の基準について, 薬食安発 0328007 号. 2005 年 3 月 28 日.
23) Chhikara P, Hammad TA. Rethinking drug safety signal detection and causality assessment in the age of AI: the risks of incomplete data and biased insights. Frontiers in Drug Safety and Regulation. 2025;5:1678074.
24) Nagar A, Gobburu J, Chakravarty A. Artificial intelligence in pharmacovigilance: advancing drug safety monitoring and regulatory integration. Therapeutic Advances in Drug Safety. 2025;16:20420986251361435.
25) Seinen TM, Kors JA, van Mulligen EM, Rijnbeek PR. Using structured codes and free-text notes to measure information complementarity in electronic health records: feasibility and validation study. Journal of Medical Internet Research. 2025;27:e66910.
26) Golder S, Xu D, O'Connor K, Wang Y, Batra M, Hernandez GG. Leveraging natural language processing and machine learning methods for adverse drug event detection in electronic health/medical records: a scoping review. Drug safety. 2025;48(4):321–37.
27) Sohn S, Wang Y, Wi C-I, Krusemark EA, Ryu E, Ali MH, et al. Clinical documentation variations and NLP system portability: a case study in asthma birth cohorts across institutions. Journal of the American Medical Informatics Association. 2018;25(3):353–9.
28) Moon S, McInnes B, Melton GB. Challenges and practical approaches with word sense disambiguation of acronyms and abbreviations in the clinical domain. Healthcare informatics research. 2015;21(1):35–42.
29) Usui M, Aramaki E, Iwao T, Wakamiya S, Sakamoto T, Mochizuki M. Extraction and standardization of patient complaints from electronic medication histories for pharmacovigilance: natural language processing analysis in Japanese. JMIR medical informatics. 2018;6(3):e11021.
30) Yamanouchi Y, Nakamura T, Ikeda T, Usuku K. An alternative application of natural language processing to express a characteristic feature of diseases in Japanese medical records. Methods of Information in Medicine. 2023;62(03/04):110–8.

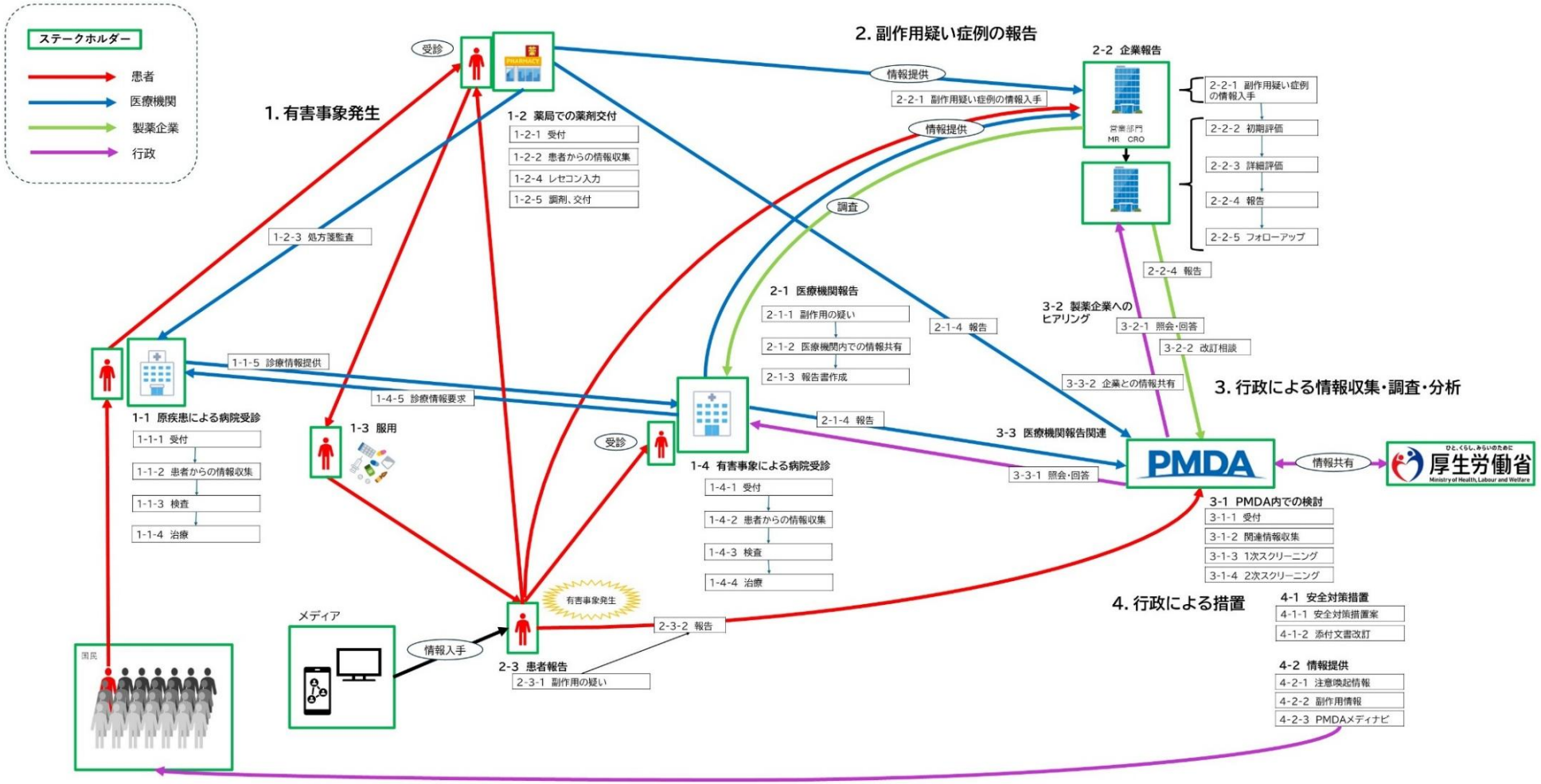

図１．PV プロセス・ハイレベルマップ
ステークホルダーとして，患者，医療機関（病院，薬局），製薬企業，行政機関（PMDA，厚生労働省），メディアを設定した．医薬品と副作用が疑われる事象の因果関係評価に必要な情報には，事象が発生する以前のものが含まれると考え，当該医薬品が処方されるに至った「原疾患による病院受診」を起点とした．

表 1. 各ユースケースにおけるリスク層と対策に関する考え方（案）

ユースケース（PV/ICSR 中心）	AI の関与（例）	リスク層（目安）[*1]	最小安全構成（共通・土台）	追加統制（高リスク時）
① 受付・取り込み（RPA[*2]/ETL[*3]、形式変換）	収集・転記・整形	低	閉域／非学習・非保持／HITL[*4]＋証跡	（原則不要）※データ結合や自由記載を扱うなら入手力(I/O)制御・再識別対策
② 自由記載（narrative 等）からの項目抽出	副作用名/日付/併用薬など抽出候補	中	同上	影響評価（誤抽出の影響）＋根拠紐付け（どこから抽出したか）＋変更管理
③ MedDRA コーディング候補提示	PT[*5]/LLT[*6] 候補提示	中	同上	フェイルセーフ（不確実は人手へ）＋根拠紐付け＋継続監視
④ 重篤性/報告要否の一次スクリーニング支援	ルール＋LLM[*7] で「要確認」提示	高	同上	影響評価＋フェイルセーフ（偽陰性対策）＋根拠紐付け＋変更管理・監視
⑤ 症例トリアージ（優先度付け）	緊急度スコア、担当振分け	中～高	同上	影響評価（遅延の影響）＋フェイルセーフ（上位取りこぼし防止）＋継続監視
⑥ 既知/未知判定の支援（添付文書等との照合）	記載有無・一致度提示	高	同上	影響評価＋根拠紐付け（どの版・どの記載に基づくか）＋変更管理（添付文書改訂に追随）
⑦ 因果関係評価の補助（一次評価案）	考え方の提示/根拠候補	高	同上	影響評価＋フェイルセーフ（提案は提案に限定、判断は人）＋根拠紐付け＋監視
⑧ フォローアップ質問案の自動生成	追跡質問の草案	中	同上	自由記載 I/O 制御（個人情報・過剰収集抑制）＋変更管理
⑨ 定型文書（症例サマリ/報告書草案）生成	下書き生成	中～高	同上	根拠紐付け（引用元）＋I/O 制御＋変更管理（プロンプト/テンプレ改訂管理）
⑩ シグナル検出（候補抽出～優先度）	解析結果の解釈・クラスター提示	高	同上	影響評価＋フェイルセーフ（見逃し対策）＋継続監視（性能/ドリフト）＋変更管理
⑪ 添付文書改訂“案”の作成支援	文案/論点整理	高	同上	影響評価＋根拠紐付け（裏付け・版）＋変更管理・監視
⑫ 監査・査察向けの記録自動整形	監査証跡・一覧化	低～中	同上	（原則不要）※ただし監査証跡の完全性確保＝根拠紐付け/変更管理が有益

[*1] リスク層は, AI の関与が意思決定に与える影響の程度で低・中・高に層別化し, 高リスクでは追加統制を検討, という考え方に沿って分類した. [*2] RPA: Robotic Process Automation (定型業務の自動化), *3 ETL: Extract, Transform, Load (データの抽出・変換・格納) , [*4] HITL: Human-in-the Loop (人間が関与・確認する仕組み), [*5] PT: Preferred Term (MedDRA における基本語), [*6] LLT: Lowest Level Term (MedDRA における下層語), [*7] LLM: Large Language Model (大規模言語モデル).