



Title	有害事象自発報告データベースの活用による医薬品安全性研究の推進
Author(s)	浅野, 弘斗
Citation	大阪大学, 2026, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/105583
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

博士学位論文

有害事象自発報告データベースの 活用による医薬品安全性研究の推進

医療薬学分野

浅野 弘斗

目次

略語	1
緒論	2
本論	4
第 1 章 FAERS を用いたパーキンソン病に対するモノアミン酸化酵素阻害薬の安全性比較 ...	4
第 1 節 背景	4
第 2 節 方法	6
第 3 節 結果	10
第 4 節 考察	18
第 5 節 小括	21
第 2 章 シグナル検出プラットフォーム AEST の開発	22
第 1 節 背景	22
第 2 節 方法	23
第 3 節 結果	28
第 4 節 考察	34
第 5 節 小括	36
第 3 章 ローカル LLM の性能評価と特化型 LLM 開発の可能性：JNEP を用いた評価	37
第 1 節 背景	37
第 2 節 方法	39
第 3 節 結果	47
第 4 節 考察	53
第 5 節 小括	55
結論	56
謝辞	57
参考文献	58
主論文	67

略語

本論文において以下に示す略語及び略号を用いた。

AE: Adverse Event

AEST: Adverse Events Signal-detection Tool

ADR: Adverse Drug Reaction

FAERS: FDA Adverse Event Reporting System

FDA: Food and Drug Administration

HDF5: Hierarchical Data Format version 5

HLGT: High Level Group Term

HLT: High Level Term

ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use

JNEP: Japanese National Examination for Pharmacists

LLM: Large Language Model

LLT: Lowest Level Term

LoRA: Low-Rank Adaptation

MAO: Monoamine oxidase

MAO-A: Monoamine oxidase A

MAO-B: Monoamine oxidase B

MedDRA: Medical Dictionary for Regulatory Activities

PD: Parkinson's Disease

PRR: Proportional Reporting Ratio

PT: Preferred Term

RBD: REM Sleep Behavior Disorder

ROR: Reporting Odds Ratio

SMQ: Standardised MedDRA Queries

SOC: System Organ Class

VDS: Virtual Dataset

緒論

副作用（ADR: Adverse Drug Reaction）は、患者の健康状態を悪化させ、入院の増加や医療費の上昇を招く主要な要因の 1 つである¹。特に高齢者では、多剤併用や薬物代謝能の低下といった要因により ADR が発生しやすく、高齢者の入院のうち 3.3-23.1%が ADR に起因すると報告されている²。また、敗血症や好中球減少症などの重篤な ADR は致命的となる場合もある³。そのため、ADR の発生抑制に向けた取り組みや研究は公衆衛生および医療経済の両面から重要である。多くの ADR は、有害事象（AE: Adverse Event）の発生を契機として把握される。AE は、医薬品との因果関係を問わず生じた全ての好ましくない、あるいは意図しない疾病またはその徴候を指す⁴。AE の中から ADR を推定することが、安全性評価の基盤となる。

ファーマコビジランスは、臨床試験および市販後の調査を通じて AE の情報を収集・把握し評価する取り組みとして位置づけられている⁵。臨床試験は医薬品と AE の因果関係を評価しやすい利点を持つ⁶一方で、被験者数や試験期間の制約から網羅的な安全性評価には限界がある。臨床試験の限界を補完するために、市販後も継続して安全性を評価する体制が医薬品規制当局（規制当局）を中心に整備されてきた⁷。

市販後の安全性評価において、FDA などの規制当局が運営する自発報告システムは中心的な役割を果たしている。医療従事者、製薬企業、患者などが経験または把握した AE は規制当局に報告され、システムに蓄積される。自発報告システムの代表例として、米国食品医薬品局（FDA: Food and Drug Administration）が運用する FAERS（FDA Adverse Event Reporting System）が挙げられる⁸。FAERS は 2025 年 10 月時点で 3,000 万件を超える世界最大規模の自発報告データを有し⁹、市販後安全性調査における主要なデータベースの一つである。収集された報告は FAERS などのデータベースを通じて公開されており、1960 年代以降、医薬品安全性情報の主要な供給源となっている¹⁰。近年では、AE の報告件数が増加傾向にあり、特に患者からの自発的な報告が増えている¹¹。患者が自身の AE を共有することは、他の患者の安全な薬物治療を支援するとともに、新たな ADR の発見にもつながる^{12,13}。FAERS をはじめとする自発報告データベースから ADR を推定する主要な統計手法として、シグナル検出が用いられている。これは、特定の医薬品と AE の報告頻度が統計的に高い組み合わせを抽出する方法であり、検出されたシグナルは ADR 候補とみなされる¹⁴。

最近の安全性研究では、同効薬間の安全性比較が注目を集めており、臨床における治療薬選択の客観的根拠となることが期待されている^{15,16}。しかし、依然としていくつかの課題が残されている。第一に、シグナル検出に化学構造や薬理学的特性に基づく医学・薬学的知見を統合することで、ADR の機序への理解が深まる¹⁷ものの、これらの知見を活用した安全性比較は未だ十分とは言えない。第二に、FAERS をはじめとする自発報告データベースには、医薬品名などの表記の

ゆれや重複報告などデータベース設計に関する課題が存在し、解析結果の信頼性を損なう要因となりうる。安全性比較において信頼性の高い比較結果を得るためには、医薬品名の正規化や重複排除といったデータ品質の向上が不可欠である¹⁶。また、FAERSのような膨大なデータを効率的に処理・可視化するためには、大規模データ解析を支える計算基盤の整備も求められる¹⁸。第三に、シグナル検出結果を臨床的に理解するためには、統計学・薬理学・化学などの専門知識を統合的に活用する必要がある、多分野連携の難しさがFAERSの活用における障壁となっている^{17,19,20}。以上の課題に対し、本研究では以下の3点を研究課題として設定した。

第1章では、化学構造・薬理学的特性とADR発現の関連を解析する探索的な研究として、モノアミン酸化酵素B (MAO-B: Monoamine oxidase B) 阻害薬の安全性比較を実施した。MAO-B阻害薬はパーキンソン病治療薬として広く使用されてきたが、アルツハイマー病²¹や中枢神経損傷²²に対する有効性も報告されるなど適応拡大が期待される薬効群であり、治療薬選択に加え新薬開発への有用な知見となりうる。

第2章では、FAERSのデータ品質向上と計算基盤構築の課題に対応するため、Webアプリケーション「AEST (Adverse Events Signal-detection Tool)」を開発した。AESTは、標準化されたFAERSを基盤とし、高速かつ再現性の高いシグナル検出の実施と結果の可視化を実現する解析環境を提供する。多様な専門的背景を有するユーザーが直感的に解析を実行できる設計とし、安全性研究の裾野拡大を目指す。

第3章では、AESTで得られた統計解析結果の臨床的理解と活用を支援する機能の将来的な実装に向け、大規模言語モデル (LLM: Large Language Model) の応用可能性を検討した。特に、データ機密性および運用安定性 (外部システムに依存しない) という観点から、オンプレミス環境で動作するLLM (ローカル LLM) に焦点を当てた。ADRの機序的理解には薬理学をはじめとする多角的な視点が必要とされ、広範な薬学的知識がファーマコビジランスを支える基盤となる²⁰。そこで、薬学教育における知識・技能を総合的に評価する薬剤師国家試験を用いて、LLMの薬学領域における性能評価と専門知識付与のためのファインチューニング手法を、探索的に検証した。利用者ニーズに応じて安全性に関する情報を提供するLLMの開発に向けた基礎的検討である。

以上の研究を通じて、医療現場での意思決定を支援するエビデンスの創出と、自発報告システムの利活用基盤の拡充を図り、医薬品安全性研究のさらなる発展に寄与しうる新たな知見を得たので、ここに博士論文としてまとめた次第である。

本論

第1章 FAERS を用いたパーキンソン病に対するモノアミン酸化酵素阻害薬の安全性比較

第1節 背景

緒論で述べた通り、同一の薬理作用を有する医薬品であっても化学構造や代謝経路の違いにより、発現する ADR の種類や頻度が異なる可能性がある。本章では、一例として、パーキンソン病 (PD: Parkinson's Disease) の治療に用いられるモノアミン酸化酵素 B (MAO-B: Monoamine oxidase B) 阻害薬に着目し、FAERS を用いた安全性比較を実施した。

Global Burden of Disease 2019 による推計では、PD の有病率は 1990 年から 2019 年にかけて一貫して増加しており、2019 年には全世界の罹患者数が 850 万人を超えるとされる²³。PD は主にドパミンの枯渇による神経伝達異常を原因とし、筋固縮、無動、姿勢保持障害といった運動症状に加え、非運動症状も引き起こす。治療では、症状に応じてドパミン補充療法、ドパミン作動薬、COMT 阻害薬、MAO-B 阻害薬などを組み合わせ、症状を適切に管理することが求められる²⁴。特に MAO-B 阻害薬は、近年アルツハイマー病²¹や中枢神経系損傷²²への有効性も探求されており、新薬開発が活発に進められている領域である。

脳や消化管に高発現する酵素であるモノアミン酸化酵素 (MAO: Monoamine oxidase) には、MAO-A (Monoamine oxidase A) と MAO-B の 2 つのアイソフォームが存在する。線条体で MAO-B を選択的に阻害するとドパミン濃度が上昇し抗 PD 作用を示す。一方、消化管等で非選択的に MAO の阻害が生じた場合、チラミンの分解が阻害され、高血圧等の重篤なチラミン反応を引き起こす可能性がある。そのため、MAO-B 阻害薬の開発では高い選択性が重要である。現在、セレギリン、ラサギリン、サフィナミドは、いずれもパーキンソン病治療に用いられる MAO-B 阻害薬である。これらは単剤、あるいはレボドパ製剤との併用で、レボドパの作用時間延長、レボドパ治療開始時期の遅延といった効果をもたらす²⁵。一方、構造や薬理作用には MAO-B 阻害薬間で特徴的な違いが見られる (Figure 1)。

セレギリンは不可逆的な MAO-B 阻害薬で、1989 年に FDA から PD 治療薬として承認された。低用量では脳内の MAO-B へ選択的に作用するが、高用量では血中濃度の上昇に伴う選択性の低下が生じる。また、セレギリンの化学構造はアンフェタミン骨格を含み、初回通過効果によって L-アンフェタミンおよび L-メタンフェタミンへと代謝される²⁶。主な副作用として頭痛、めまい、不眠、悪心などが挙げられる²⁷。代謝物が引き起こすアンフェタミン様の副作用、特に心血管系や中枢神経系への影響については、研究間で見解が分かれている。複数の研究では心血管系の AE との関連が示唆されている²⁸⁻³³。一方で、通常用量での関連を否定する報告もあり^{34,35}、さらなる検証が必要である。

ラサギリンも不可逆的かつ選択的な MAO-B 阻害薬であり³⁶、2006 年に FDA の承認を受けた。アンフェタミン骨格を持たず、セレギリンより高い MAO-B 選択性を持つことから、副作用が少ないと期待されている。実際に、セレギリンからラサギリンへ切り替えることで睡眠障害が改善したという報告もある³⁷。さらに臨床試験の結果に基づき、ラサギリンではチラミン含有食品の摂取制限が不要となった³⁸。一方、ラサギリンの添付文書には依然としてチラミン反応への注意喚起が残っており³⁹、さらなる調査の必要がある。

サフィナミドは α -アミノアミド誘導体で、近年承認された新しい MAO-B 阻害薬である。セレギリンやラサギリンとは異なり、可逆的に MAO-B を阻害する点が特徴である。MAO-B への選択性もさらに高いため、安全性の向上が期待されている。加えて、サフィナミドはドパミン作動性作用だけでなく、電位依存性ナトリウムチャネルおよびカルシウムチャネルの阻害やグルタミン酸放出抑制といった、非ドパミン性の作用も併せ持つ^{40,41}。多面的な作用が、他の MAO-B 阻害薬との安全性の違いを生む可能性がある。

MAO-B 阻害薬は PD 治療で重要な役割を担うが、医薬品ごとに選択性、化学構造、非ドパミン性作用の有無などが異なり、ADR に影響を及ぼす可能性がある。これまで臨床試験やメタアナリシスによる安全性評価は行われてきたが、チラミン反応やアンフェタミン効果のような、臨床における重大な安全性の課題を捉えるには自発報告データベースなどを用いた網羅的な解析が不可欠である。そこで本章では、FAERS を用いてセレギリン、ラサギリン、サフィナミドを対象に安全性比較を実施し、治療薬選択に資する情報を提供することを目的とする。

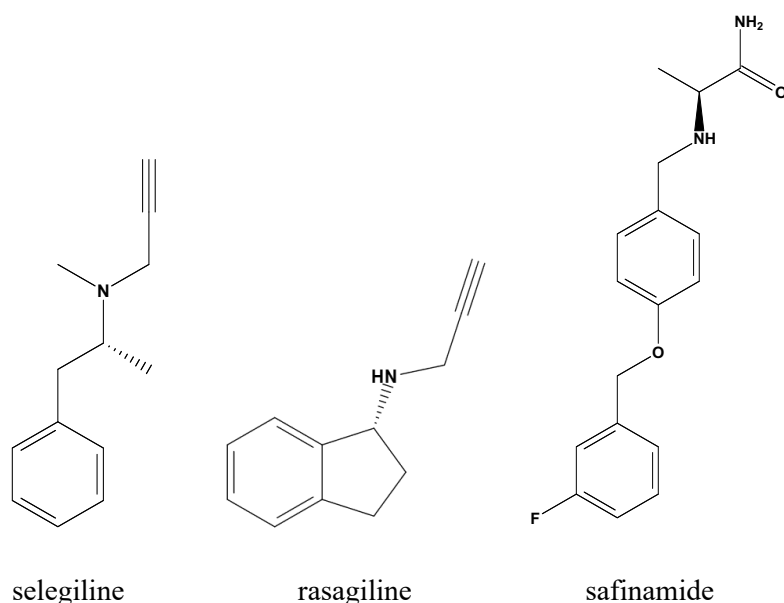


Figure 1. Chemical structures of MAO-B inhibitors

第2節 方法

データソース

FAERS は、専門家、消費者、製薬企業から自発報告を収集しており、四半期ごとに DEMO、DRUG、INDI、OUTC、REAC、RPSR、THER の 7 つのテーブル形式で公開されている。FAERS の公式ウェブサイト⁴²から、2004 年第 1 四半期から 2022 年第 2 四半期までのデータを取得した。

前処理

FAERS に収集されたデータには、報告のばらつき、不整合、誤記などが存在するため、データを標準化した。

DEMO テーブルは、年齢、体重、性別、発生日、報告国などの患者背景情報を含む。体重は kg に統一し、変換できないものは除外した。年齢も同様に年単位に変換し、変換不能なものは解析対象外とした。発生日は YYYY-MM-DD 形式で記載された日付のみを使用した。報告国は ISO 3166⁴³に準拠して標準化し、誤記は手動で修正した。

DRUG テーブルは、AE 報告時に使用された医薬品名を含む。医薬品名には、一般名、製品名、成分名、略称など多様な記載が存在するため、KEGG DRUG⁴⁴⁻⁴⁷を用いて医薬品名を統一した。まず KEGG ID をキーとした医薬品名辞書 (KEGG DICT) を作成し、FAERS の医薬品名に KEGG ID を割り当てた。配合剤の場合は成分ごとに KEGG ID を付与した。医薬品名が見つからない場合は成分名 (prod_ai) を使用した。さらに、FAERS で新たに確認された医薬品名で KEGG DICT を更新し、再度割り当てを試みた。最終的に割り当てできなかった医薬品名はそのまま使用した。以上の操作と手動での確認作業を通じて、セレギリン、ラサギリン、サフィナミドについて表記のゆれを統一した。

REAC テーブルに含まれる AE は、MedDRA (Medical Dictionary for Regulatory Activities)⁴⁸に基づき記載される。MedDRA は国際医薬品規制調和会議 (ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use) により策定された国際標準医学用語集であり、MedDRA の著作権および MedDRA®の登録商標は ICH が有する。器官別大分類 (SOC: System Organ Class)、高位グループ語 (HLGT: High Level Group Term)、高位語 (HLT: High Level Term)、基本語 (PT: Preferred Term)、下層語 (LLT: Lowest Level Term) の 5 階層構造を有する。FAERS に報告される AE は PT の階層から選択される。報告時に使用されていた MedDRA バージョンの違いによる AE 表記のばらつきを解消するため、全ての AE 表記を MedDRA version 25.0 に基づいて統一した。

THER テーブルは投薬情報を含み、YYYY-MM-DD 形式で記載された開始日・終了日のみを解析に使用した。

INDI テーブルは投薬理由が記載される。PT に準拠する記載は MedDRA ver 25.0 の PT に統一し、その他の記録は除外した。

テーブル統合およびデータ集計

データの統合に関し、DRUG、THER、INDI テーブルについては primaryid と drug_seq をキーとして、その他のテーブルについては primaryid をキーとして統合した。医薬品名、開始日、終了日、適用症、AE、性別、AE 発生日、体重、年齢、報告国が同一の記録を重複とし、最後の報告を除く全ての記録を削除した。さらに、医薬品開始日が AE 発生日より後の報告は除外した。

データの集計では、年齢・性別・報告国を含む背景情報を集計した。男性・女性以外の性別は集計から除外した。報告国は、UN M49⁴⁹ に基づいて地域ごとに集計した。不明な記録は集計から除外した。年齢については平均値と標準偏差 (SD) を算出し、性別と報告国については割合を算出した。

シグナル検出

シグナルは、ROR (Reporting Odds Ratio)⁵⁰ により検出した。ROR はオランダの Lareb など規制当局でも用いられる、代表的なシグナル検出の指標の 1 つである。ROR の 95%信頼区間下限および報告数 (n11) がともに 1 を超えた場合をシグナルと定義した (Figure 2)⁵¹。また、セレギリンは剤形により適応が異なり、経口剤はパーキンソン病治療、パッチ製剤 (EMSAM) はうつ病治療に用いられる。本研究では適応の違いを考慮し、パッチ製剤を使用した症例、または投薬理由がうつ病関連である症例を除外した。さらに、検出されたシグナルは Primary SOC に従い SOC に変換し集計した。

	Suspected ADR	Other ADRs
Suspected drug	n11	n12
Other drug	n21	n22

Where

- n11 = Number of suspected ADR requested when suspected drug is used
- n12 = Number of other ADRs requested when suspected drug is used
- n21 = Number of suspected ADR requested when other drugs are used
- n22 = Number of other ADRs requested when other drugs are used

$$ROR = \frac{n11/n12}{n21/n22} = \frac{n11 \times n22}{n12 \times n21}$$

$$95\%CI = e^{\log(ROR) \pm 1.96 \sqrt{\frac{1}{n11} + \frac{1}{n12} + \frac{1}{n21} + \frac{1}{n22}}}$$

Figure 2. Calculation Methods for ROR

アンフェタミン様作用の検証

セレギリンに関しては、アンフェタミン様作用の発現が懸念されてきたが、その定義は明確ではない。そこで、本研究ではアンフェタミンのシグナル検出結果との類似度と、添付文書の警告事項との重複度という 2 つの指標を用いて、アンフェタミン様作用を定量的に評価した。具体的には、以下の 2 つの方法を用いた。

第一に、アンフェタミンと MAO-B 阻害薬（セレギリン、ラサギリン、サフィナミド）それぞれについて検出されたシグナル間の類似度を、Dice 係数および Simpson 係数（式 1、2）を用いて比較した。これらの係数は類似度計算に用いられる指標である⁵²。

式 1

$$Dice(X,Y) = \frac{2 \times |X \cap Y|}{|X| + |Y|}$$

式 2

$$Simpson(X,Y) = \frac{|X \cap Y|}{\min(|X|, |Y|)}$$

X はアンフェタミンに関するシグナルの集合を表し、Y は各 MAO-B 阻害薬（セレギリン、ラサギリン、またはサフィナミド）のシグナルの集合を表す。係数値が大きいほど、両者間の類似性が高いことを示す。アンフェタミンのシグナル検出結果と各 MAO-B 阻害薬（セレギリン、ラサギリン、サフィナミド）のシグナル検出結果との類似度を比較することで、医薬品ごとのアンフェタミン様作用を評価した。さらに、オッズ比が高いシグナルは特に関連性が強いと考えられるため、アンフェタミンについて検出されたシグナルのうち、オッズ比が上位 50%、25%、10% のシグナルを用いた感度分析を実施した。また、アンフェタミンにおいて特に懸念される SOC（精神、神経、心臓、および血管）に属する PT についても比較した。

第二に、アンフェタミンの添付文書に記載された警告事項と各 MAO-B 阻害薬で検出されたシグナルの重複度を確認した。警告事項は重篤な心血管イベント、精神障害、成長抑制、けいれん、末梢血管障害（レイノー現象含む）、セロトニン症候群、視覚障害を含む。これらの警告に該当する MedDRA および SMQ（Standardised MedDRA Queries）と、MAO-B 阻害薬のシグナルを比較した。

第3節 結果

FAERS（2004 年第 1 四半期 - 2022 年第 2 四半期）のうち、セレギリン 3,784 件、ラサギリン 10,584 件、サフィナミド 1,007 件の報告を解析対象とした（Table 1）。平均年齢はそれぞれ 68 歳、70 歳、71 歳であり、女性の割合は 40～44%と、MAO-B 阻害薬間で顕著な差は認められなかった。報告症例の性別については、いずれも男性が女性を上回った。報告地域は、いずれの医薬品も北米が最も多かったが、それ以外の地域では分布に差がみられた。セレギリンは東アジア、特に日本での報告割合が高く、東アジアからの報告 991 件のうち、950 件を日本が占めた。

Table 1. Baseline information

	Selegiline	Rasagiline	Safinamide
Age in years (mean $\pm$ SD (n))	67.8 $\pm$ 12.7 (2667)	70.0 $\pm$ 10.7 (7209)	71.4 $\pm$ 9.8 (713)
Number of combination drugs ^a (mean $\pm$ SD)	8.5 $\pm$ 6.2	7.5 $\pm$ 5.6	7.8 $\pm$ 4.9
Monotherapy ^b (n, %)	34 (0.8%)	883 (8.3%)	57 (5.6%)
Female (n, %)	1517 (42.6%)	3998 (40.5%)	415 (44.3%)
Area (n, %)			
Northern America	1523 (45.2%)	6082 (59.6%)	377 (40.4%)
Western Europe	201 (6.0%)	1497 (14.7%)	251 (26.9%)
Eastern Asia	991 (29.4%)	708 (6.9%)	29 (3.1%)
Northern Europe	310 (9.2%)	799 (7.8%)	62 (6.6%)
Southern Europe	231 (6.9%)	439 (4.3%)	197 (21.1%)
Eastern Europe	23 (0.7%)	169 (1.7%)	0 (0.0%)
South America	42 (1.2%)	94 (0.9%)	16 (1.7%)
Western Asia	4 (0.1%)	130 (1.3%)	0 (0.0%)
Central America	8 (0.2%)	110 (1.1%)	0 (0.0%)
Australia and New Zealand	22 (0.7%)	90 (0.9%)	1 (0.1%)
Southern Asia	5 (0.1%)	40 (0.4%)	1 (0.1%)
South-eastern Asia	11 (0.3%)	19 (0.2%)	0 (0.0%)
Southern Africa	0 (0.0%)	19 (0.2%)	0 (0.0%)
Caribbean	0 (0.0%)	5 (0.0%)	0 (0.0%)
Northern Africa	0 (0.0%)	1 (0.0%)	0 (0.0%)
Micronesia	0 (0.0%)	1 (0.0%)	0 (0.0%)

In this table, unreported data were excluded from the calculation of both the count (n) and the percentage (%).

a. The number of combination drugs was counted when other drugs were reported for the same primary ID.

b. Monotherapy is counted when no other drugs are reported for one primary ID.

MAO-B 阻害薬におけるシグナルの検出

セレギリン、ラサギリン、サフィナミドにおいて検出されたシグナルは、それぞれ 585、672、254 件の PT であった。

全てに共通するシグナルは 111 件あり、SOC ごとに集計した (Table 2)。神経系障害 (n=30, 27.0%)、精神障害 (n=27, 24.3%)、損傷、中毒および処置合併症 (n=18, 16.2%) が多く、全体の 65%以上を占めた。神経疾患においては、オン・オフ現象 (セレギリン OR 37、95% CI 29-46；ラサギリン OR 45、95% CI 40-51；サフィナミド OR 78、95% CI 58-104) など、パーキンソンズ様のシグナルが主に検出された。精神障害に関しては、衝動制御障害 (セレギリン OR 92、95% CI 71-120；ラサギリン OR 72、95% CI 60-87；サフィナミド OR 12、95% CI 3-46)、性欲亢進症 (セレギリン OR 145、95% CI 115-184；ラサギリン OR 100、95% CI 84-120；サフィナミド OR 73、95% CI 39-136) などのシグナルが検出された。損傷、中毒および処置合併症に関しては、運動機能に関連する転倒や骨折、損傷のほか、ストーマ関連のシグナルも 8 件検出され、PD に対する経腸療法に起因する可能性がある。

一方、医薬品ごとの特異的なシグナルとして、セレギリン 308 件、ラサギリン 367 件、サフィナミド 87 件の PT が検出された (Table 3)。セレギリンでは肝胆道系障害、ラサギリンでは腫瘍および血管障害がそれぞれ特異的なシグナルとして認められた。一方、サフィナミドに特有の SOC はみられなかったが、眼障害に関するシグナルは他の 2 つの医薬品より多かった。

Table 2. Common Signals (SOC level) and the number of preferred terms (PT).

SOC	Number of PT
Nervous system disorders	30 (27.0%)
Psychiatric disorders	27 (24.3%)
Injury, poisoning and procedural complications	18 (16.2%)
General disorders and administration site conditions	10 (9.0%)
Product issues	7 (6.3%)
Respiratory, thoracic and mediastinal disorders	3 (2.7%)
Musculoskeletal and connective tissue disorders	3 (2.7%)
Gastrointestinal disorders	3 (2.7%)
Surgical and medical procedures	2 (1.8%)
Vascular disorders	1 (0.9%)
Social circumstances	1 (0.9%)
Skin and subcutaneous tissue disorders	1 (0.9%)
Renal and urinary disorders	1 (0.9%)
Metabolism and nutrition disorders	1 (0.9%)
Investigations	1 (0.9%)
Infections and infestations	1 (0.9%)
Eye disorders	1 (0.9%)
Total	111

Table 3. Product-specific signals (SOC level) and the number of preferred terms (PT).

SOC	Selegiline	Rasagiline	Safinamide
Investigations	37 (12.0%)	22 (6.0%)	6 (6.9%)
Nervous system disorders	35 (11.4%)	49 (13.4%)	16 (18.4%)
Cardiac disorders	32 (10.4%)	7 (1.9%)	2 (2.3%)
Injury, poisoning and procedural complications	30 (9.7%)	43 (11.7%)	10 (11.5%)
Gastrointestinal disorders	23 (7.5%)	21 (5.7%)	0 (0.0%)
Psychiatric disorders	22 (7.1%)	52 (14.2%)	3 (3.4%)
Respiratory, thoracic and mediastinal disorders	16 (5.2%)	12 (3.3%)	6 (6.9%)
Infections and infestations	14 (4.5%)	12 (3.3%)	7 (8.0%)
General disorders and administration site conditions	11 (3.6%)	42 (11.4%)	6 (6.9%)
Surgical and medical procedures	11 (3.6%)	12 (3.3%)	4 (4.6%)
Musculoskeletal and connective tissue disorders	9 (2.9%)	6 (1.6%)	3 (3.4%)
Social circumstances	8 (2.6%)	5 (1.4%)	2 (2.3%)
Neoplasms benign, malignant and unspecified (include cysts and polyps)	7 (2.3%)	24 (6.5%)	1 (1.1%)
Renal and urinary disorders	7 (2.3%)	5 (1.4%)	2 (2.3%)
Product issues	6 (1.9%)	10 (2.7%)	1 (1.1%)
Skin and subcutaneous tissue disorders	6 (1.9%)	8 (2.2%)	2 (2.3%)
Metabolism and nutrition disorders	6 (1.9%)	7 (1.9%)	2 (2.3%)
Eye disorders	5 (1.6%)	4 (1.1%)	8 (9.2%)
Reproductive system and breast disorders	5 (1.6%)	4 (1.1%)	0 (0.0%)
Blood and lymphatic system disorders	5 (1.6%)	3 (0.8%)	1 (1.1%)
Hepatobiliary disorders	5 (1.6%)	0 (0.0%)	0 (0.0%)
Vascular disorders	2 (0.6%)	12 (3.3%)	3 (3.4%)
Congenital, familial and genetic disorders	2 (0.6%)	2 (0.5%)	0 (0.0%)
Endocrine disorders	2 (0.6%)	1 (0.3%)	1 (1.1%)
Immune system disorders	1 (0.3%)	2 (0.5%)	0 (0.0%)
Pregnancy, puerperium and perinatal conditions	1 (0.3%)	1 (0.3%)	0 (0.0%)
Ear and labyrinth disorders	0 (0.0%)	1 (0.3%)	1 (1.1%)
Total	308	367	87

チラミン反応の検証

チラミン反応のシグナルは、ラサギリン（OR 626, 95%CI 192–2038）でのみ検出され、セレギリンおよびサフィナミドでは検出されなかった。ただし、抗うつ薬として使用されるセレギリンのパッチ製剤（EMSAM）では、チラミン反応の報告が4件あった。

サフィナミドの非ドパミン性作用

夢内容の行動化により外傷などを伴うレム睡眠行動障害（RBD: REM Sleep Behavior Disorder）はPD患者の約半数に認められる⁵³。サフィナミドの非ドパミン性作用がRBDに有効である可能性が報告されており⁵⁴、セレギリン、ラサギリン、サフィナミドにおいてRBDがシグナルとして検出されるか否かに着目した。シグナルとしてRBDが検出されたのはセレギリン（OR 23, 95%CI 7–70）とラサギリン（OR 51, 95%CI 31–82）のみであり、サフィナミドでは検出されなかった。

アンフェタミン様作用の検証

アンフェタミンと各MAO-B阻害薬のシグナル類似度を示すDice係数およびSimpson係数をTable 4に示す。いずれの比較において類似度は0.61以下であった。セレギリンとアンフェタミンの間に、他のMAO-B阻害薬と比較してシグナルの明確な共通性は認められなかった。

次に、アンフェタミンの添付文書⁵⁵に記載された警告と、MAO-B阻害薬において検出されたシグナルとの重複を確認した（Table 5）。セレギリンと他のMAO-B阻害薬との間に明確な差異は確認されなかった。

Table 4. The Dice's and Simpson's coefficients of AE signals were detected between amphetamine and each MAO-B inhibitor

	Coefficients	Amphetamine	Selegiline	Rasagiline	Safinamide
Total signals		1530	585	672	254
All signals	D ^a	Ref.	0.18	0.23	0.12
	S ^b	Ref.	0.33	0.37	0.43
Top 50% in odds ratio	D	Ref.	0.10	0.11	0.05
	S	Ref.	0.11	0.11	0.10
Top 25% in odds ratio	D	Ref.	0.05	0.06	0.03
	S	Ref.	0.07	0.08	0.04
Top 10% in odds ratio	D	Ref.	0.02	0.02	0.00
	S	Ref.	0.05	0.05	0.01
PTs included in SOC (psychiatric, neurological, cardiac, and vascular disorders)	D	Ref.	0.36	0.43	0.22
	S	Ref.	0.48	0.54	0.60
PTs included in SOC (psychiatric, neurological)	D	Ref.	0.40	0.47	0.23
	S	Ref.	0.58	0.59	0.61

^aD, Dice's coefficient; ^bS, Simpson's coefficient.

Table 5. The headlines and subheadlines listed as WARNINGS in an amphetamine package insert and the number of the corresponding MedDRA terms in each product signal.

Headline	Subheadline	The number of corresponding MedDRA terms (PT)	Selegiline	Rasagiline	Safinamide
Serious cardiovascular events	Sudden deaths	3	1	0	0
	Stroke	11	0	2	0
	Myocardial infarction	41	1	2	1
	Hypertension	90	8	8	2
Psychiatric adverse events	Pre-existing psychosis worse	1	1	1	1
	Bipolar illness & Emergence of new psychotic or manic symptoms	7	2	2	1
	Aggression	1	1	1	1
Long-term suppression of growth		3	0	0	0
Seizures		91	3	2	3
Peripheral vasculopathy, including Raynaud's phenomenon		63	0	2	2
Serotonin syndrome		1	1	1	1
Visual disturbance		96	2	4	5

第4節 考察

PD 治療薬である 3 種の MAO-B 阻害薬（セレギリン、ラサギリン、サフィナミド）に対し、FAERS を用いてシグナル検出を実施した。先行研究では、PD の発症率は 60 歳以降に急増し、男性が女性よりも高く、人種・民族による違いも報告されている⁵⁶。FAERS の報告は PD の発症率そのものを直接示すものではないが、報告者の背景情報は疫学研究で報告されている PD 患者の特徴と同様の傾向を示した。併用薬数は平均 8 剤程度と多剤併用の傾向がみられたが、これは PD が身体的・精神的併存症を伴うことに関連すると考えられる⁵⁷。一方、国・地域ごとの報告頻度は、医薬品ごとに異なっていた。要因の一つに各国での承認年の違いが挙げられる。日本では 2007 年から 2018 年までセレギリンのみが承認されていたのに対し、ラサギリンとサフィナミドは欧州（それぞれ 2005 年、2015 年）、米国（2006 年、2017 年）、日本（2018 年、2019 年）と、特にラサギリンの承認時期が地域間で異なる。このため、東アジア、特に日本でのセレギリンの報告頻度が高いと考えられる。

肝胆道系障害のシグナルは、特にセレギリンで検出された。地域別では、日本からの報告が特に多くみられた。例えば、肝嚢胞に関する報告は全 10 件であり、そのうち 8 件が日本からの報告であった。セレギリンは高濃度で MAO-B 選択性が低下し、MAO-A 阻害作用が現れる⁵⁸。セロトニンの代謝酵素である MAO-A の活性低下は、肝硬変など肝疾患に関与するセロトニンの蓄積につながりうる^{59,60}。特に、セレギリンの代謝酵素である CYP2C19 や CYP2B6 に多型がある場合や薬物相互作用によっては、セレギリンの血中濃度上昇が生じやすくなる^{61,62}。CYP2C19 の Poor Metabolizer は、白人で約 5%、日本人で約 20%とされる⁶³。さらに、日本人 PD 患者を対象としたセレギリンの臨床試験では、100 件中 2 件で肝胆道系の副作用が報告されている⁶⁴。これらの先行研究を考慮すると、セレギリンによる肝胆道系 ADR の可能性、特に日本人での発生頻度の上昇を示唆している。今後の更なる検証は必要だが、肝胆道系障害には注意が求められる。

血管障害および腫瘍のシグナルは、特にラサギリンで検出された。血管障害に関しては、ラサギリンが血圧に影響を及ぼすことが複数の臨床試験で報告されている。早期パーキンソン病患者を対象にラサギリン単剤療法の有効性を評価した TEMPO 試験では、2 mg 投与群で臥位収縮期血圧の有意な上昇が認められた⁶⁵。Stern らの報告では、2 および 4 mg/日のラサギリン投与で血圧がわずかに低下する傾向がみられた⁶⁶。レボドパ治療中で、1 日 2.5 時間以上の OFF 時間を伴う重篤な運動変動を有する進行期パーキンソン病患者を対象にラサギリンの補助療法の有効性を評価した PRESTO 試験では、進行期 PD 患者においてラサギリン 0.5 mg/日投与で立位時の収縮期・拡張期血圧が低下した⁶⁷。レボドパ治療中で運動変動を有する進行期パーキンソン病患者を対象としラサギリンの補助療法の有効性を評価した LARGO 試験では、1 mg/日投与で起立性低血圧の発現率が 2%と報告されている⁶⁸。これらの臨床試験の結果は、ラサギリンが血管系障害、特に血圧に影響を及ぼす可能性を示しており、本結果を支持する。

腫瘍に関しては、TEMPO 試験で 2 mg 群 132 例中 3 例が新たに悪性腫瘍（悪性黒色腫、前立腺癌、扁平上皮癌）と診断された⁶⁵。TEMPO 試験第 2 相では 371 例中 5 例が新たに悪性腫瘍（大腸癌、皮膚扁平上皮癌 2 例、基底細胞癌、黒色腫）と診断された⁶⁹。ラサギリンの早期開始が未治療パーキンソン病患者の疾患進行を抑制するか評価した ADAGIO 試験では、ラサギリン 1 mg/日投与患者 1 例が投与 72 週で黒色腫を発症した⁷⁰。PD と腫瘍の関係については、PD は黒色腫と正の相関、他の腫瘍とは負の相関が報告されている⁷¹。近年、ラサギリンが黒色腫のリスク因子となる可能性も指摘されている⁷²。本研究でも、黒色腫関連シグナルはラサギリンのみで検出されており、リスク因子である可能性を支持する。他のリスク因子（生活習慣、基礎疾患、報告バイアスなど）の影響も考慮する必要はあるが、ラサギリン使用時は腫瘍関連 AE、特に黒色腫に注意が必要である。

眼障害と関連するシグナルは、サフィナミドで多く検出された。ラットを用いた動物実験では薬剤誘発性網膜萎縮が報告されており、ヒトでの使用前に眼科的検査が必要とされる。第 3 相臨床試験でも、サフィナミドの使用に伴う視覚障害や白内障が報告されている⁷³。報告数が少なく背景因子も不明ではあるが、サフィナミド使用時の眼障害にも注意が必要である。

チラミン反応に関して、臨床試験ではラサギリン単剤投与群（n=38）ではチラミン反応は認められなかったが、レボドパ併用群 0.5 mg/日投与患者 22 例中 3 例でチラミン反応が生じた⁷⁴。本検討でも、ラサギリンのシグナルとしてチラミン反応が低頻度（n=3）で検出された。2004 年以降の FAERS で、チラミン反応は 37 例報告され、ラサギリンでの報告数は約 8%を占める。一方、セレギリンではチラミン反応のシグナルは検出されなかったが、先行研究ではセレギリン（5 mg, 1 日 2 回）のチラミン感受性係数はラサギリン（1 mg/日）より高かったと報告されている³⁸。FAERS においても、EMSAM（セレギリンパッチ製剤）とチラミン反応の組み合わせが 4 件あり、

ラサギリンおよびセレギリン使用時にはチラミン反応に注意が必要である。

サフィナミドのグルタミン酸放出抑制作用に着目した最近の研究では、REM 睡眠行動障害に対する有効性が報告されている⁵⁴。PD 患者における RBD の有病率は 46～58%とされる⁵³。RBD は REM 睡眠中に激しい運動を示す疾患であり、グルタミン酸、グリシン、GABA がその発症に重要な役割を果たす⁷⁵。本検討でも RBD に着目したところ、サフィナミドでは RBD のシグナルは検出されず、非ドパミン性作用により RBD の発現リスクが低い、あるいは抑制的に働く可能性が示唆された。ただし、サフィナミドは報告数が限られているため、検出力不足によりシグナルが検出されなかった可能性も考慮すべきである。イタリアの調査では、2017 年に MAO-B 阻害薬を使用した患者 1,059 例中、セレギリン 502 例 (47%)、ラサギリン 161 例 (15%)、サフィナミド 396 例 (37%) であった⁷⁶。サフィナミドは比較的新しく、今後の処方数増加に伴い FAERS 報告数も増加すると予測される。そのため、サフィナミドの RBD 抑制効果については継続的な評価が必要である。

アンフェタミンの使用により、興奮、幻覚、自殺企図、胸痛などが生じることが報告されている⁷⁷⁻⁷⁹。また、心脳血管性の死亡例も報告されている⁸⁰。セレギリンの代謝物である L-アンフェタミンは、D-アンフェタミンの約 1/10 の薬理活性とされるが、線条体では活性が同等程度であると報告されている⁸¹。本章の研究では、MAO-B 阻害薬におけるアンフェタミン様作用について、シグナルの類似度と添付文書記載事項との比較により検証した。その結果、セレギリンのアンフェタミン様シグナルは他の MAO-B 阻害薬と同程度であり、セレギリンに明確なアンフェタミン様作用は認められなかった。この結果は、セレギリンの臨床試験でアンフェタミン様 AE が認められていないこと⁶⁴や、リスザルおよびラットを用いた研究でも通常用量でアンフェタミン様作用が認められなかったこととも一致する^{34,35}。

本章の研究は FAERS に基づいており、いくつかの限界がある。まず、FAERS の精査を行ったものの、報告バイアスなど FAERS 固有の制約を完全に排除することはできない⁸²。また、併用薬や投与期間に関する情報が不十分なため、単剤使用者を正確に区別することは困難である。加えて、シグナルは因果関係を示すものではなく、あくまでリスクの可能性を示唆するに留まる。そのため、今後は、電子カルテやレセプトデータなどのリアルワールドデータを活用し、MAO-B 阻害薬の安全性について更なる検証を行うことが重要である。

第5節 小括

同効薬間の安全性比較は、臨床における治療薬選択に客観的な根拠を与え、適正な薬物治療の推進に資する。特に、シグナル検出に化学構造や薬理・薬物動態といった薬学的知見を統合することは、ADRの機序的理解を深めるうえで有用である。

本章では、パーキンソン病治療に用いられる MAO-B 阻害薬（セレギリン、ラサギリン、サフィナミド）に着目し、FAERS データを用いたシグナル検出結果の比較を行った。まず、セレギリンの化学構造に起因するとされるアンフェタミン様作用への懸念について、Dice 係数および Simpson 係数を用いたシグナルの類似度に基づき評価した。その結果、セレギリン特異的なアンフェタミン様作用は確認されなかった。一方で、CYP 代謝との関連が疑われる肝胆道系障害のシグナルが認められた。次に、ラサギリンでは、MAO-B 選択性に関連するチラミン反応がシグナルとして検出された。また、非ドパミン性作用を有するサフィナミドでは、PD 患者の約半数に認められる REM 睡眠行動障害はシグナルとして検出されなかった。このように、シグナル検出結果を化学構造・薬理・動態と関連付けて考察した結果、臨床的懸念を評価する上で有用な知見が得られた。

一方、FAERS に代表される自発報告データベースを用いた研究では、解析手法やデータ前処理のばらつきが大きく、研究間の結果比較を困難にしている⁸³。臨床現場での客観的な治療薬選択を支援するためには、研究者や医療者がそれぞれの目的に合わせ、再現性の確保された環境下で FAERS データを柔軟に活用できる解析プラットフォームの整備が望まれる。

第2章 シグナル検出プラットフォーム AEST の開発

第1節 背景

第1章では、FAERS データを用いたシグナル検出を通じて MAO-B 阻害薬の安全性を解析した。同一の薬理作用を有する医薬品であっても、化学構造や代謝経路の違いが発現する ADR の差異に寄与し得ることを明らかにした。この知見は、臨床における治療薬選択の根拠となりうる。近年、このような同効薬間の安全性比較は注目を集めており、てんかん治療薬やドパミン作動薬を対象とした報告もされている^{84,85}。しかし、これらの報告の多くは依然として小規模に留まり、最大でも6種類程度の医薬品比較に限定されている。さらなる課題として、報告間での FAERS 前処理方法や解析期間が統一されていないことが挙げられる。解析条件の違いは、研究間での結果の整合的な解釈を困難にし、治療薬選択の根拠としての有用性低下につながりうる。

加えて、計算基盤の不足も大きな障壁となっている。FAERS は 3,000 万件を超える報告を含む巨大データベースである⁹。FAERS の解析には、プログラミングスキルや計算資源、統計解析ソフトウェアの習熟などが求められる。そのため、臨床の知見を有する医療従事者や他分野の研究者が実際に FAERS を活用・解析するにはハードルが高い¹⁸。安全性評価の質を高めるためには多角的な視点からの検証が不可欠であり、実現には各分野の専門家が FAERS を活用できる解析プラットフォームの整備が必要である。FAERS を対象とした解析サービスやツールは有用であるものの、2025 年時点ではサービスの停止や機能面での制約があり、また一部の商用ツールは有償である⁸⁶⁻⁸⁹。

本章では、上記の課題に包括的に対応するため、新たな Web アプリケーション「AEST (Adverse Events Signal-detection Tool)」を開発した。AEST は3つの設計思想に基づいて構築されている。第一に、標準化されたデータ基盤の提供である。統一的な前処理プロトコルを適用した FAERS を基盤とし、研究間での結果の比較可能性と再現性を担保する。第二に、高速なシグナル検出環境の実現である。分散並列処理を導入し、数秒単位での高速解析を実現する。第三に、アクセシビリティの最大化である。研究者から医療従事者に至るまで、多様な専門的背景を持つユーザーに対し、シグナル検出から可視化までを一貫して実行できる直感的なインターフェースを無償で提供する。

以上により、医療従事者や研究者など、多様な立場からの FAERS を活用した安全性評価研究の実施を可能とし、医薬品安全性研究の裾野拡大とエビデンス創出の加速に貢献する。

第2節 方法

データソース

FAERS の公式ウェブサイトより 2004 年第 1 四半期から 2025 年第 2 四半期までの利用可能なデータ全てを取得し、解析対象とした⁴²。FAERS の前処理は第 1 章と同様の手順に基づいて実施したが、データ拡張およびアプリケーション構築に対応するため、下記の処理内容を再設計した。

MedDRA version の統一

報告時に使用されていた MedDRA バージョンの違いによる AE 表記のばらつきを解消するため、全ての AE 表記を MedDRA version 28.1 に基づいて統一した。

年齢の範囲設定

ヒトの生理学的限界を考慮し、年齢が 0～125 歳の範囲外の値は欠損値として扱った⁹⁰。

医薬品名の正規化

DRUG テーブルでは医薬品名が自由記述形式で登録されており、一般名、商品名、有効成分名、略称等の多様な表記が混在している。表記ゆれはシグナル検出の計算に必要な集計数を変動させ、解析精度の低下要因となり得るため、医薬品名の正規化処理を実施した。なお、医薬品名にアルファベットを含まない場合は正規化処理から除外した。

DiAna Dictionary を基盤として AEST Dictionary を新規に構築した。DiAna Dictionary は、2004 年から 2022 年 12 月までの FAERS において全医薬品エントリの 98.94%、固有の医薬品名 (unique drug entries) の 43.72% をカバーする医薬品名正規化辞書である⁹¹。DiAna Dictionary は公式ウェブサイトより取得した 2025 年 6 月 25 日更新版を使用した⁹²。AEST Dictionary の構築は二段階の工程で実施した。第一段階では、DiAna Dictionary を活用して FAERS 医薬品名の正規化を行った。第二段階では、同辞書で未対応である FAERS の医薬品名について、正規表現による追加処理を実施した。非英字記号を区切りとして複合記載や付加情報を分割し、分割して得られた成分候補を既存成分と照合した。一致した成分は新たに AEST Dictionary へ登録し、カバー率の向上を図った。

AEST データの構築

FAERS の高速処理を実現するため、まず文字列データの数値化を実施した。医薬品名や AE 名などの文字列項目に一意の整数 ID (0 : 欠損値、1 以降 : 有効データ) を割り当て、全データを数値配列に変換した。ID と文字列の対応辞書は別ファイルで保存し、解析結果を表示する際の逆変換に使用した。数値化したデータは四半期単位で分割し、HDF5 形式で保存した。ファイル内では医薬品名や AE などの項目ごとに階層構造を構築した。HDF5 形式は、ファイル全体をメモリに展開せず必要な箇所を局所的に読み込むことを可能とし、メモリ効率とアクセス速度を向上させる⁹³。四半期ごとに分割されたファイルを統合的に処理するため、HDF5 Virtual Dataset (VDS) を導入した⁹⁴。VDS の導入により、複数の物理ファイルを仮想的に単一データセットとして参照できるようになり、個別ファイルの結合処理を省略した全期間にわたる横断的な検索・解析を実現した。AEST 構築までの一連の開発プロセスを Figure 3 に示した。

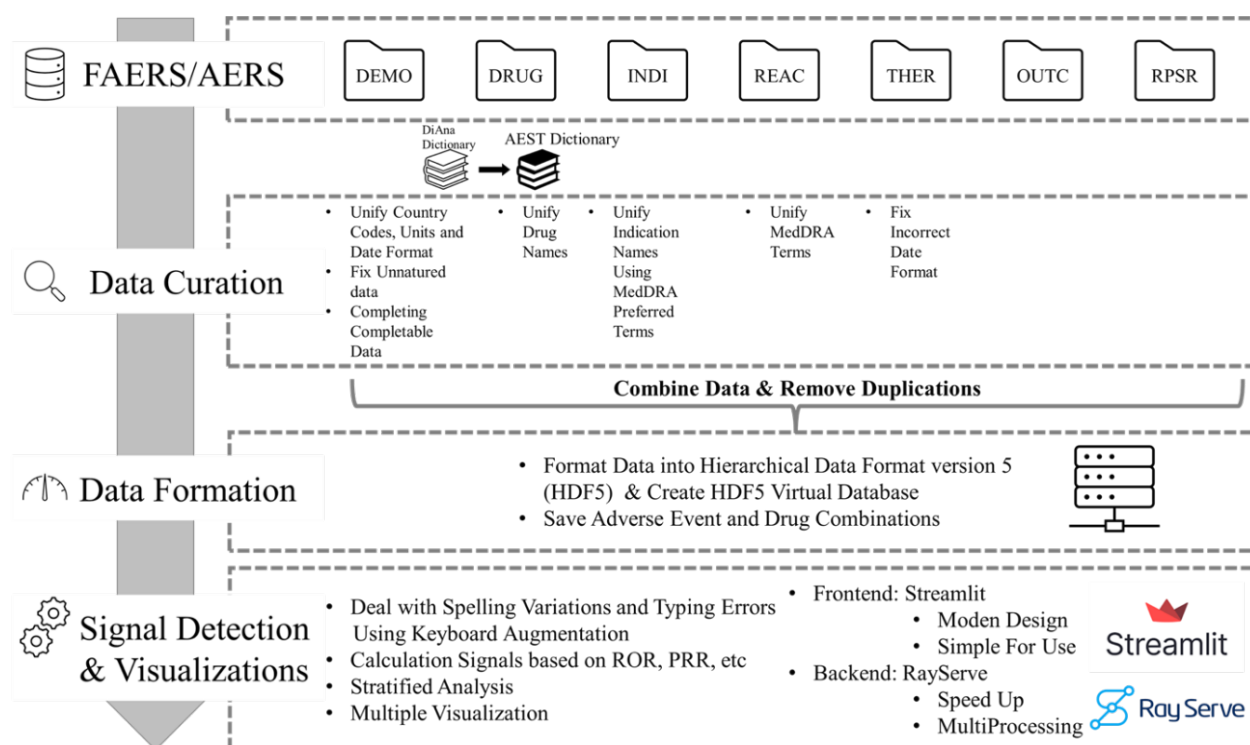


Figure 3. Sequential development process for the construction of AEST

AEST システム開発

技術概要

AEST は Python を用いて構築した。画面表示には Streamlit、データ処理には Ray を採用した。Streamlit は、Web アプリケーションを構築できる Python ライブラリである⁹⁵。Ray は大規模データの並列処理を効率的に実行できる分散処理ライブラリである⁹⁶。ユーザーが画面で医薬品名や解析条件を入力すると、サーバー上で Ray が FAERS の集計を並列実行する (Figure 4)。その後、シグナル検出を実施し、結果は自動的に画面に表やグラフとして表示される。

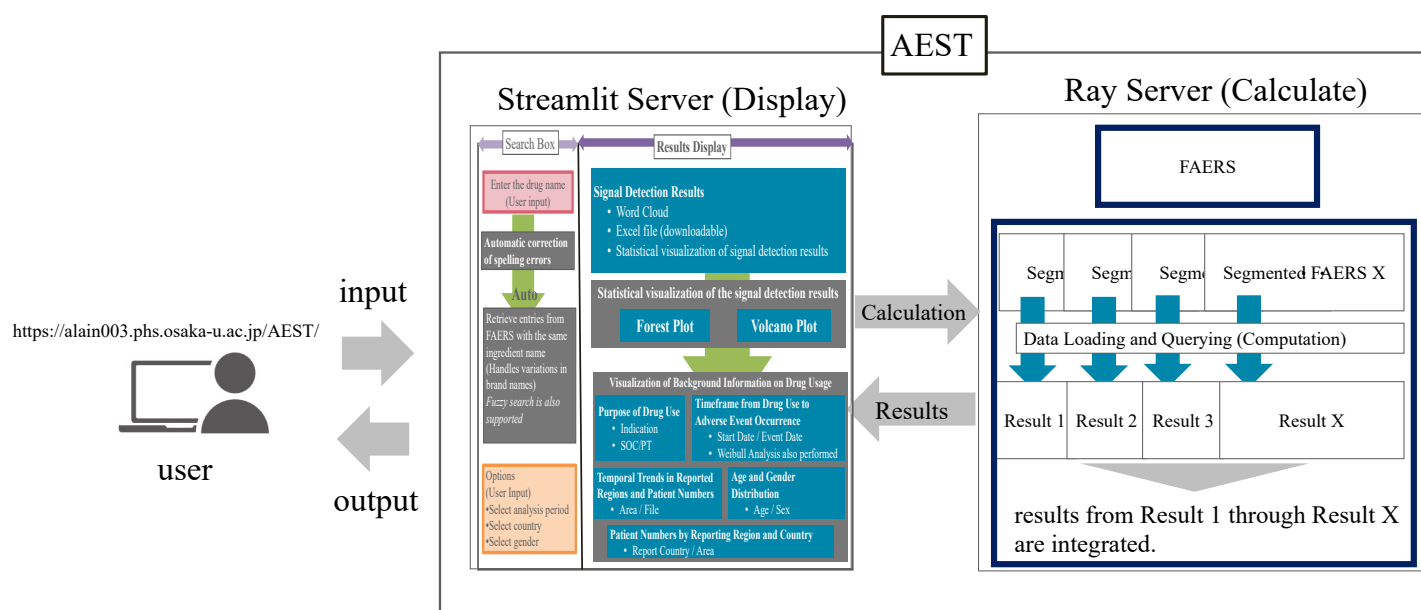


Figure 4. Distributed signal detection workflow in AEST

User inputs are processed via a Streamlit server for display and automatic correction, while calculations are parallelized over segmented FAERS data on the Ray server, with integrated results visually summarized.

シグナル検出計算の高速化

シグナル検出では、報告件数を n11 (対象医薬品と対象 AE の件数)、n12 (対象医薬品と非対象 AE の件数)、n21 (非対象医薬品と対象 AE の件数)、n22 (非対象医薬品と非対象 AE の件数) に分類した 2x2 分割表を基に、PRR (Proportional Reporting Ratio) と ROR を Table 6 の基準に沿って算出する。なお、AEST では 1 人の患者が同一医薬品に対して複数の AE を報告した場合、医薬品と AE のペアを個々に独立した報告として計上し、対応する分割表の各セルに反映させる。本検討では PRR または ROR で基準を満たしたものを、シグナルとする^{50,51,97}。

また、計算効率を高めるため、AEST では FAERS 全体における各 AE の総報告数をあらかじめ集計し、解析時に必要な報告数を参照する仕組みとしている。解析のプロセスでは、まず対象医薬品に関連する各 AE の報告数 (n11) を合算することで、その医薬品における全 AE の報告数 (n1+) を算出する。個々の AE を解析する際には、この n1+ から該当 AE の n11 を減算することで、対象医薬品に関連するその他の AE の報告数 (n12) を求める。AEST のデータベースでは、FAERS データを医薬品名と PT の組み合わせ単位で格納しているため、対象医薬品でフィルタリングするだけで、該当 AE の報告数 (n11) および非対象医薬品における当該 AE の報告数 (n21) を迅速に算出できる。ただし、解析期間、国、性別などの条件を指定してシグナル検出を行う場合には、これらの条件を満たすデータを動的に抽出し、その部分集合に対して AE ごとの報告数を再集計した上で解析を行う。

Table 6. Calculation Methods for ROR and PRR

Method	Formula	Signal Threshold
ROR (Reporting Odds Ratio)	$ROR = (n11 \times n22) / (n12 \times n21)$	
	95% CI	$n11 \geq 2$
	$= \exp(\ln(ROR))$	lower limit of 95% CI of $ROR > 1$
	$\pm 1.96 \times \sqrt{1/n11 + 1/n12 + 1/n21 + 1/n22}$	
PRR (Proportional Reporting Ratio)		$n11 \geq 3$
	$PRR = (n11 / (n11 + n12)) / (n21 / (n21 + n22))$	$PRR \geq 2$
		$\chi^2 \geq 4$

分散並列処理システム

AEST では、FAERS の解析における計算コストの課題に対応するため、分散並列処理フレームワーク Ray を導入した。具体的には、FAERS を複数の小さなデータ単位に分割し、それぞれを独立したタスクとして非同期に並列集計するシステムを実装した。本処理環境には、20 vCPU および 30 GiB のメモリを割り当て、限られたリソース下でのスループット向上を図った。

システムが多数のユーザーからの同時アクセスに耐えられるかを検証するため、負荷試験を実施した。AEST で解析可能な 5,721 成分について、2004 年第 1 四半期から 2025 年第 2 四半期までの FAERS を用い、シグナル検出および背景情報の可視化が完了するまでの応答時間（秒）と応答が得られないエラーの発生率を測定した。検索リクエストの送信には、Playwright によるブラウザ自動操作スクリプトを用い、並列処理数（スレッド数）を 1、6、12、24 と段階的に増加させた。各条件下で、検索から結果表示までの処理時間を計測し、同時アクセス時の性能を評価した。

医薬品名検索機能

AEST では、ユーザー入力に基づいて適切な医薬品名を推定する検索機能を実装し、利便性の向上を図った。医薬品名には成分名、商品名、略称など多様な表記が存在し、入力時のスペルミスも考慮すると、完全一致のみに依存する単純な文字列照合では十分な検索精度を確保できない⁹⁸。そこで、AEST ではユーザー入力文字列と AEST 辞書に登録された医薬品名の類似度を Jaro-Winkler 距離を用いて算出し、最もスコアが高い候補を返却する方式を採用した。Jaro-Winkler 距離は文字列先頭の一致度を重視する特性を持ち、文字列のスペルミスに対して頑健であるため⁹⁹、入力誤りを含む場合でも適切な医薬品名を推定することが期待される。加えて、本機能はユーザー入力とシステム内部で管理される医薬品名との対応付けを行う役割も担う。

この検索機能の頑健性を定量的に評価するため、タイピングエラーを含む検索語に対して正しい医薬品名を特定できるかを検証した。AEST でシグナル検出が可能な全 5,721 成分名の文字列に対し、nlpaug ライブラリ¹⁰⁰の KeyboardAug 機能を用いて、QWERTY 配列に基づくタイピングエラーを擬似的に生成した。エラー発生率（aug_char_p）を 0%から 50%まで 5%刻みで段階的に増加させ、各エラー率における検索後の一致率（正答率）を測定した。

第3節 結果

AEST データの概要

AEST のデータセットは医薬品名と AE の組み合わせから構成され、総計 575,452,020 件のレコードを含む。四半期あたりの平均レコード数は 6,691,302 件（標準偏差 6,094,644）であり、全体として増加傾向にある（Table 7、Figure 5）。データ記載率については、医薬品名（Drug name）および AE（PT）は全報告に記載されている一方、開始日（Start date）や AE 発生日（Event date）など一部項目は記載率が相対的に低かった。また、医薬品名や AE 名などの文字列項目は一意の整数 ID へ符号化し、全データを数値配列として再構成した。AEST で使用した FAERS の元データサイズは 5.47 GiB であったが、数値化したデータを VDS および名称復元用ファイルとともに格納した結果、データサイズは 1.25 GiB（元データの約 23%）にまで圧縮された。

Table 7. Summary table of data completeness for key fields in AEST

Data total: 575,452,020 records (mean per file 6,691,302; SD $\pm$ 6,094,644)	
Filling rate (number/total)	
Drug name	100.0% (575,452,020/575,452,020)
PT	100.0% (575,452,020/575,452,020)
Indication PT	99.9% (575,451,928/575,452,020)
Start date	14.7% (84,453,152/575,452,020)
Event date	40.7% (234,136,717/575,452,020)
End date	85.0% (489,402,120/575,452,020)
Age	71.8% (413,091,651/575,452,020)
Sex	88.8% (510,802,519/575,452,020)
Reporter country	96.3% (553,877,001/575,452,020)

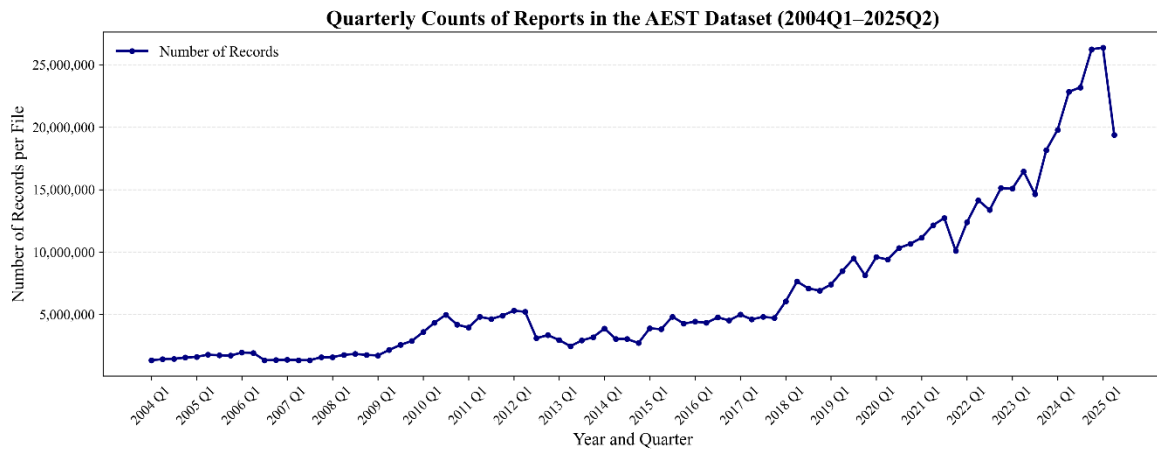


Figure 5. Quarterly trends in record counts

Quarterly counts of reports in AEST from 2004Q1 to 2025Q2, showing a continuous increase in report volume with accelerated growth in recent years.

インターフェース

AEST では、医薬品名を入力することでシグナル検出から結果の可視化まで自動的に実行される (Figure 6)。医薬品名の入力では、Jaro-Winkler 距離による曖昧検索機能を実装した。文字列の先頭一致度を重視した Jaro-Winkler 距離に基づく類似度計算により、タイピングエラーなどの入力ミスに対しても自動補正が機能する。入力後は、シグナル検出が自動実行され、各種背景情報のグラフとともに可視化された解析結果が提示される。さらに、ユーザーはインターフェース上でのマウス操作により、解析条件（期間、国、年齢層など）を柔軟に指定できる。

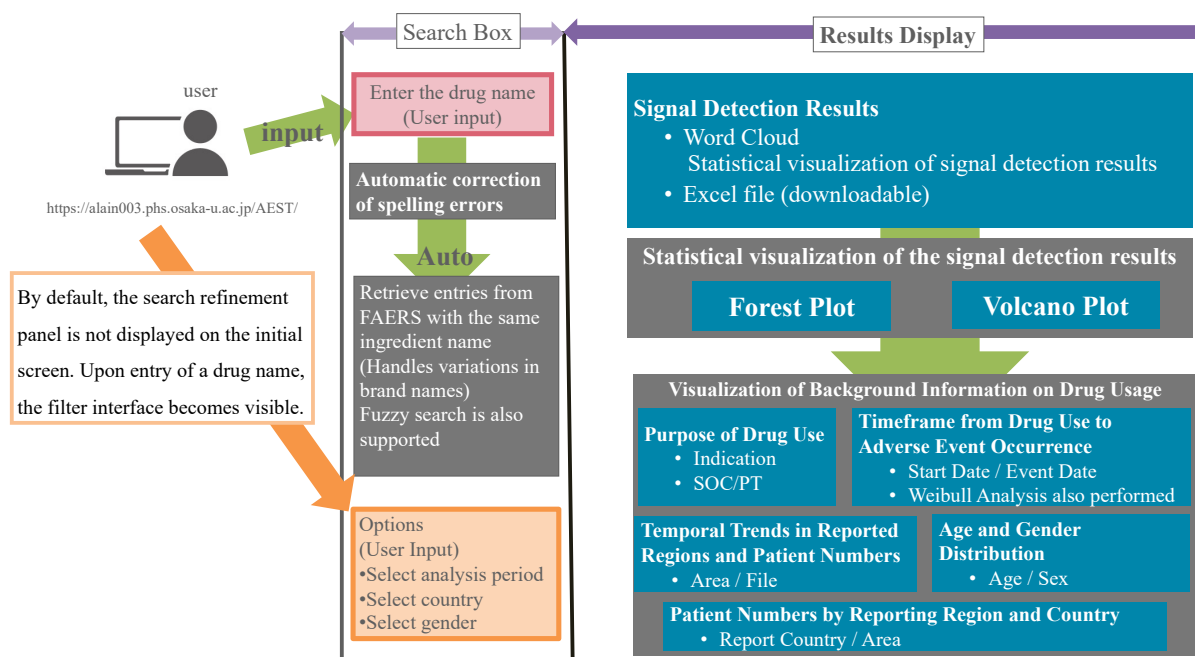


Figure 6. Overview of the AEST interface

AEST integrates automated spelling correction, brand name harmonization, and an interactive search panel triggered by drug name input. It generates word cloud, statistical plots (forest and volcano plots), and detailed summaries of drug usage patterns, time-to-onset profiles, and patient demographics to support in-depth pharmacovigilance analyses.

医薬品名標準化

DiAna Dictionary を基盤とし、正規表現などのテキスト処理を活用して拡張した AEST Dictionary を新たに作成した。2025 年 6 月 25 日更新版の DiAna Dictionary は、2004 年第 1 四半期から 2025 年第 2 四半期までの FAERS に含まれる医薬品名のうち 6,466 成分に対応している。AEST データセットに対する DiAna Dictionary 単体のカバー率は、全医薬品名に対して 97.7%、ユニークな医薬品名に対しては 39.81% (351,563/883,058) であった。これに対し、AEST Dictionary では、カバー率がそれぞれ 99.03% および 66.84% (590,277/883,058) となり、医薬品名の網羅性が向上した。

曖昧検索精度の評価

ユーザー入力誤りに対する Jaro-Winkler 距離による補正精度を、シグナル検出可能な全 5,721 成分を対象に評価した (Figure 7)。タイプミス発生確率 (aug_char_p) が 0.25 以下の条件では、平均一致率 (Matching Rate) は 90%前後と高値を維持した。一方、タイプミス発生確率が 0.3 を超えると一致率は著しく低下し、0.5 の時点では 26.1%となった。この低下は、特に入力文字列長が 6~15 文字の場合に顕著に見られた。

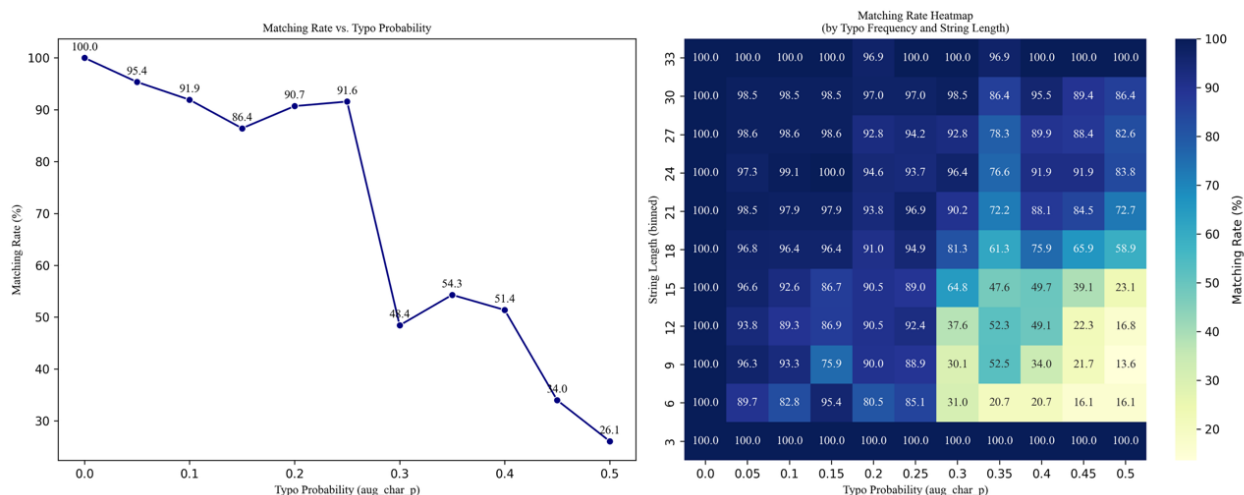


Figure 7. Matching rate according to typo probability and string length

Left: Line plot showing the decrease in matching rate as typo probability increases.

Right: Heatmap illustrating matching rates by typo frequency and string length. Matching accuracy remains high for low typo probabilities, while shorter strings and higher typo probabilities result in decreased matching performance.

システム性能とスケーラビリティ

シグナル検出に要する時間を評価するため、シグナル検出可能な全 5,721 成分を対象にシグナル検出を実施した。計算時間は、平均 6.1 秒、中央値 3.4 秒、最短 1.8 秒、最長 20.3 秒であった (Figure 8)。計算時間は、各成分の報告件数が多い場合や、成分名に加えて商品名など複数の名称表記を含む場合に長くなる傾向が認められた。

AEST では、シグナル検出対象となる 5,721 成分のうち 5,693 成分 (99.5%) で背景情報を出力できる。AEST への同時アクセス時の待機時間を評価するため、この 5,693 成分について、シグナル検出および背景情報の可視化が完了するまでの応答時間 (秒) と、アクセス負荷によるエラー発生率を測定した (Figure 9)。シングルスレッド処理では、リクエストの 95 % が完了するまでの時間 (95th パーセンタイル) は 33.56 秒、99th パーセンタイルは 76.11 秒であった。

並列処理数の増加に伴い、応答時間の延長など性能低下がみられた。6 スレッド設定では 95th パーセンタイルが 130.61 秒、エラー率は 0 % (0/5,693) であった。12 スレッド設定では 95th パーセンタイル 252.06 秒、エラー率 0.4 % (23/5,693) を示した。さらに 24 スレッドでは、95th パーセンタイルが 450.44 秒 (99th パーセンタイル : 547.88 秒) まで増加し、シングルスレッド比で約 13.6 倍に延長したほか、エラー率も 1.5 % (85/5,693) に上昇した。

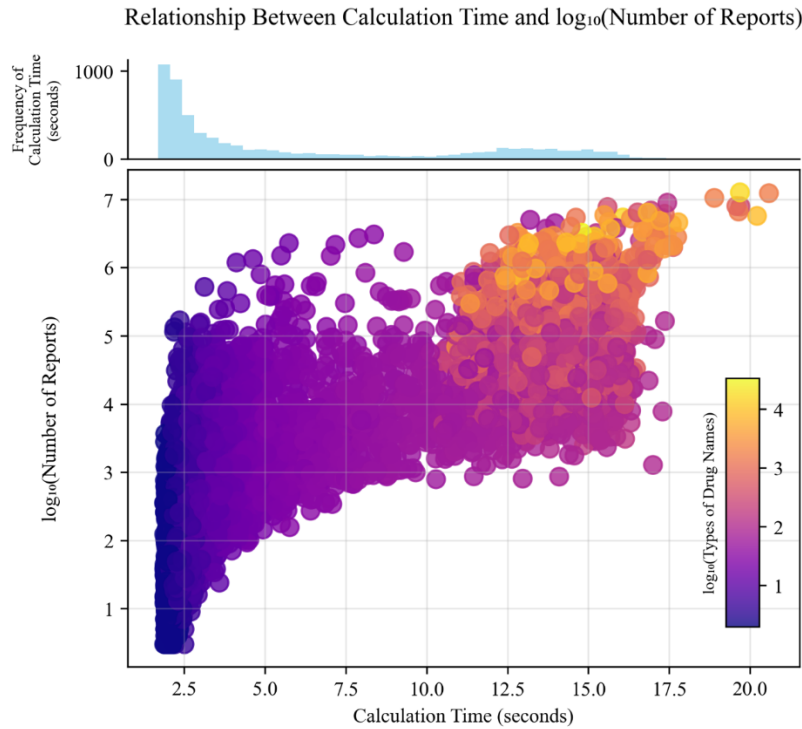


Figure 8. Relationship between calculation time and $\log_{10}$ (number of reports)

Scatter plot showing the correlation between computation time (x-axis) and the logarithmic number of reports analyzed (y-axis), with points colored by $\log_{10}$ (types of drug names). The histogram at the top displays the frequency distribution of calculation times. The plot demonstrates that computation time increases with the number of reports and that drug name diversity also influences calculation complexity.

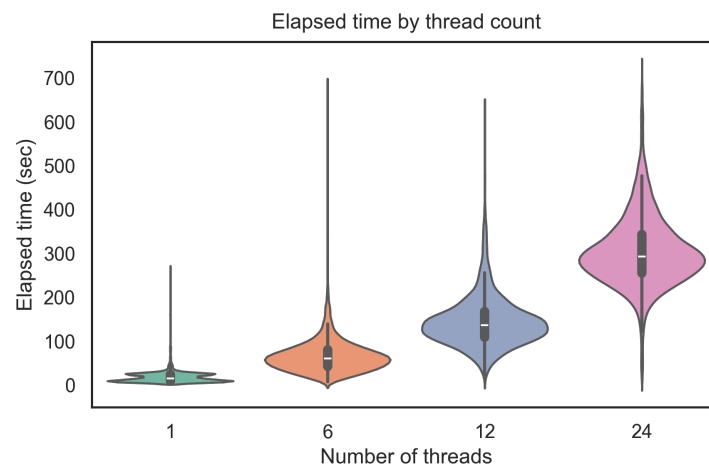


Figure 9. Distribution of elapsed time by number of threads

Violin plots displaying the distribution of elapsed processing times (seconds) for different numbers of threads (1, 6, 12, 24). The results showed that an increasing number of threads tended to result in longer processing times.

第4節 考察

FAERS を用いて医薬品のシグナルを検出できる Web アプリケーション「AEST」を開発、公開した (<https://alain003.phs.osaka-u.ac.jp/AEST>)。AEST はインターネット環境があれば、端末や OS を問わずブラウザから利用でき、専用ソフトウェアのインストールを必要としない。FAERS は四半期ごとに更新されるが、AEST は最新データが自動的に全ユーザーへ反映される設計とし、データ更新やソフトウェアのアップデート作業の負担を解消した。さらに、ユーザー登録やログインを不要とし、検索内容や入力情報を保存しない構成とすることで、利用者のプライバシーにも配慮している。

システム性能の面では、FAERS データを数値化し HDF5 形式に変換したことで、データサイズを従来の 4 分の 1 以下に削減した。また、Ray Serve による分散並列処理を採用した結果、平均 6 秒程度でシグナル検出が完了する高速処理を実現した。

シグナル検出における集計誤差や精度低下の一因となる医薬品名の多様性への対応として、DiAna Dictionary を基盤に正規表現処理を用いて拡張した AEST Dictionary を新たに構築した。これにより、全医薬品名に対するカバー率は 99.03%、重複を除いた医薬品名の種類に対するカバー率は 66.84%に達し、FAERS におけるシグナル検出の精度向上が期待される。

さらに、ユーザー利便性の観点から、Jaro-Winkler 距離に基づくあいまい検索機能を実装した。医薬品名のスペルミスを自動的に補正することで、操作性が向上すると期待される。

AEST のほかにも、主な FAERS マイニングプラットフォームとして、FDable、AERS Spider、OpenVigil、AERSMine 等があり、安全性研究などで広く活用されてきた¹⁰¹⁻¹⁰⁴。それぞれ特徴的な強みや機能を有する一方、2025 年 9 月時点ではメンテナンス停止やシグナル検出・可視化機能が不十分なものも存在する (Table 8)。これに対し AEST は、インタラクティブな Web インターフェースを介して、医薬品名の正規化やスペルミス補正といった自動処理機能に加え、期間や国などの絞り込み機能を備えた高速なシグナル検出、結果の出力、関連データの可視化といった独自の機能を提供する。

ただし、AEST には FAERS に起因する課題と技術的な限界が存在する。FAERS は報告バイアスなどの影響を受けやすく、シグナルは因果関係を直接的に示すことはできない⁸²。そのため、検出されたシグナルの臨床的解釈や応用には慎重な判断が求められる。技術面では、多数の同時アクセス時に遅延が発生する可能性がある。しかし、AEST は分散並列処理を採用しており、利用可能な計算資源に応じて動的に性能を向上できる設計となっている。本研究では、20 vCPU と 30 GiB のメモリを割り当てた環境でシステム性能を評価した。より大規模な利用環境においては、ユーザーのニーズに応じて計算資源を増強することで、同時アクセス時の遅延やエラー率を低減できると考えられる。

今後の課題として、GUI やマニュアル、チュートリアルの実装に加え、具体的な活用例を提示することにより、ユーザーによる適切な解釈および利用を支援する必要がある。また、スペルミス補正機能が解析エラーの防止にどの程度寄与したかといった定量的評価や、ユーザーからの意見を開発に反映する仕組みの導入も重要である。これらの点については、本研究の成果を基盤として、今後の実装および運用段階において段階的に改善していく予定である。

Table 8. Feature comparison between AEST and existing FAERS analysis tools.

Features	FDable	AERS Spider	OpenVigil 2.0	AERSMine	AEST (ours)
FAERS Data Mining	○	○	○	○	○
Drug Name Normalization				○	○
Handling of Drug Name Variations and Typos					○
Signal Detection with Custom Filters					○
Visualization of Signal Detection Results			○		○
Exporting Signal Detection Results			○	○	○
Visualization of Background Data					○
Operational Status (as of Sep 2025)	○		○	○	○
Pricing	\$350	free	free	free	free
URL	https://www.fda.gov/dable.com/	http://www.cdc.gov/hemoprofilin.org/AERS	https://openvigil.sourceforge.net/	https://research.cchmc.org/aers/home	https://alain03.phs.osaka-u.ac.jp/AEST

第5節 小括

第1章では、同一薬理作用を有する医薬品であっても、化学構造や代謝経路の相違がADRの発現に差異をもたらし得ることを明らかにした。この知見は、FAERSを用いたシグナル検出が、治療薬選択に対し客観的な根拠を与える手段として有用であることを示唆している。一方で、FAERS解析においては、前処理方法や解析条件が研究間で統一されておらず、得られた結果を臨床応用へと発展させる上で再現性や信頼性などの課題が残る。また、FAERSデータ解析には一定の計算資源やツールが必要であり、研究機関以外、特に臨床現場などでの活用は容易ではない。

本章では、計算資源に依存せず再現性の高い安全性解析を実施できるよう、FAERS解析支援ツールAEST (<https://alain003.phs.osaka-u.ac.jp/AEST>)を開発し、Webアプリケーションとして公開した。AESTは、直感的なユーザーインターフェースと高速な処理性能を両立している。実際、大阪大学の「学部学生による自主研究奨励事業」において、学部生が本ツールを用いて「1型糖尿病発症に関連する要因の探索」を行い、その成果を報告している¹⁰⁵。このように、本ツールは研究や教育の場においても活用可能な操作性と実用性を備えている。さらに、データセットの軽量化により、AESTはソースコードおよびデータベース一式をソースコード管理プラットフォーム上で共有可能である。ローカル環境での利用を可能とし、データ更新など将来的な継続性も確保した。今後は、研究機関に留まらず、臨床や教育を含む多様な領域での活用が期待される。

現在のAESTは実装の初期段階にあり、さらなる機能拡張や解析結果の活用方法について、今後検討の余地がある。AESTは、FAERSに収載された医薬品を網羅的かつ同一条件下で解析することを可能とする点に特徴があり、有害事象調査の効率化や適正な薬物治療の推進に寄与し得ると考えられる。一方で、解析結果をエビデンスとして整理し、規制当局や医療現場での活用につなげるためには、引き続き検討が必要である。また、多様な専門分野の研究者が有害事象データに触れる機会を広げ、臨床知見を医薬品開発へと還元していくことは重要である。そのため、AESTの利用者層の拡大や活用促進に向けた取り組みについても、今後の検討課題である。

今後の展望として、研究者や医療従事者など多様なバックグラウンドを持つユーザーの利用を想定し、インターフェースの改善による利便性の向上や、電子カルテ等のリアルワールドデータとの統合を進めることで、実用性のさらなる向上を目指す。さらに、シグナル検出結果の臨床的解釈を支援する機能を強化し、情報提供機能の充実を図ることで、医薬品安全性評価および臨床現場における意思決定支援への貢献を目指す。

第3章 ローカル LLM の性能評価と特化型 LLM 開発の可能性：JNEP を用いた評価

第1節 背景

第2章で構築した AEST は、標準化された FAERS 解析基盤上で、高速なシグナル検出と結果の可視化を実現する Web アプリケーションである。計算資源や解析環境の制約を取り払うことで、より広範なユーザー層による安全性解析への参画を可能にするものである。一方、シグナル検出結果を臨床応用するためには、多様な専門知識に基づく学際的なアプローチが求められる¹⁹。特に、AEST で検出されたシグナルの臨床的意義を医学・薬学的観点から情報提供する機能は、ユーザーの安全性研究への理解を深める上で重要である。

情報提供機能を実現する有力な技術候補として、ChatGPT のような大規模言語モデル（LLM: Large Language Model）に着目した。ファーマコビジランス活動の基盤となるのは化学・薬理学・臨床など医学・薬学的な知識であり²⁰、LLM がシグナルを適切に説明するためには、基礎から臨床に至る網羅的な専門知識を備えていることが前提となる。広範な薬学的知識を総合的に評価する日本の薬剤師国家試験（JNEP: Japanese National Examination for Pharmacists）は、薬学領域における LLM の性能を検証する有用な指標となりうる。実際に、GPT-4 は解答可能な JNEP の問題に対して、合格基準を上回る正答率を示し、日本語の医療専門知識を要する課題にも対応可能であることが示唆されている¹⁰⁶。この知見は、LLM が AEST の結果説明機能として応用できる可能性を示唆している。

しかし、ChatGPT や Gemini などインターネット経由でサービス提供者のサーバーに接続して利用するクラウド型 LLM を AEST に組み込み、医療機関内システムで運用した場合、重大なリスクが生じる可能性がある。第一に、データセキュリティの問題である。クラウド型 LLM はインターネットを介して利用者の質問文を外部サーバーに送信するため、患者情報や機密性の高い医療データが漏洩するリスクが常に存在する。患者プライバシーの保護が最優先される臨床現場において、データ漏洩リスクは LLM 導入の最大の障壁となる¹⁰⁷⁻¹¹⁰。第二に、システムの継続への懸念である。クラウド型 LLM サービス終了事例に見られるように、商用 LLM の提供停止や強制的なモデル・バージョンの更新はシステム全体の持続可能性を脅かす¹¹¹。特定用途に最適化されたシステムにとって安定的なバージョン管理は不可欠であり、クラウド型 LLM への依存は大きなリスクとなりうる。

解決策の1つとして、オンプレミス環境で動作する LLM（ローカル LLM）が注目されている。ローカル LLM はオフライン環境で継続的かつ安定的に動作可能であり、データを外部に送信しないためプライバシーを保護できる^{107,112-114}。

近年、DeepSeekをはじめとする一部のローカル LLM は、GPT シリーズに匹敵する性能を有することが報告されているが、医療・薬学分野におけるその実用的な評価や応用への取り組みは限定的である。Zong らの調査によれば、医療系国家試験を対象とした 709 件の研究のうち、ローカル LLM を採用した事例はわずか 3 件（0.4%）に留まっている¹¹⁵。とりわけ、LoRA（Low-Rank Adaptation）¹¹⁶ などのパラメータ効率的なファインチューニング技術の登場により、計算資源の制約を受けずに領域特化型モデルを構築する環境は整いつつある。しかし、薬学領域に特化したローカル LLM の性能評価や、効率的な学習戦略に関する知見は依然として十分に蓄積されていない。加えて、日本語環境において、これらの技術が薬学的専門知識の習得にどの程度有効であるか、およびどのようなファインチューニング戦略が最適であるかについても、十分な検証はなされていない。

本章では、ファーマコビジランスの基盤となる薬理学的・臨床的知識の総合的な理解を評価する試験である JNEP をベンチマークとして用い、4 つのローカル LLM を用いて薬学領域における性能を評価した。さらに、効率的なファインチューニング手法である LoRA を適用し、限られた計算資源下においても薬学的専門知識を付与できる可能性について基礎的な検討を行った。

本検討は、ローカル LLM の薬学領域における性能を探索的に把握するとともに、将来的な AEST との連携を見据え、利用者のニーズに応じた情報提供を行うシステム構築に向けた基盤的知見を得ることを目的とする。

第2節 方法

データソース

2012 年（第 97 回）から 2024 年（第 109 回）までに実施された JNEP を使用した。JNEP は、日本の厚生労働省が毎年実施する国家資格試験である¹¹⁷。試験問題の本文と科目情報は SEED から取得し、解答は YakugakuQA から取得した^{118,119}。JNEP の目的は、薬剤師として実務を行うために必要な基礎知識と理解力を評価することである。試験では、薬学知識を実際の場面に応用する能力、関連法規の理解、倫理的配慮の遵守など、臨床実践で不可欠となる幅広い能力が評価される。毎年の試験は、物理、化学、生物、衛生、薬理、薬剤、病態/薬物治療、法規/制度/倫理、実務の 9 科目にわたる 345 問の多肢選択式問題で構成される。問題は、必須問題、理論問題、実践問題という 3 つの区分からなる。

必須問題（90 問）：基礎知識に焦点を当て、通常は単一の正解形式で出題される。

理論問題（105 問）：薬学の理論に基づく問題が出題される。

実践問題（150 問）：具体的な症例などを用いて、特に薬物治療や患者管理における実践的能力を評価する。

JNEP は、単一または複数正解の多肢選択問題から構成され、指定された正解と完全に一致した場合に得点が与えられる。一部の問題は連続したセットとして出題されるため、解答には冒頭の文脈や先行問題の解答を踏まえた判断が求められる。また、試験後の検証において特定の問題が不適切と判断された場合には、問題の無効化や特定の選択肢を選んだ全受験者への得点付与など、厚生労働省による是正措置が講じられることがある。試験実施後には、全ての試験問題と正答が厚生労働省から公表される。合格基準は、受験者の得点分布に基づく相対基準により毎年設定される。第 109 回薬剤師国家試験では、全問題の得点が 420 点以上（210 問相当）、必須問題について全問題配点の 70%以上かつ各科目 30%以上、さらに倫理的・医療安全的観点から設定される禁忌肢問題の選択数が 2 問以下であることが合格基準とされた。なお、禁忌肢の具体的な問題番号や内容は公表されていない。

データ準備

本研究で使⽤した LLM は画像処理能力を有していないため、化学構造式や図表などの視覚情報を含む問題は、単問・連問を問わず除外した。同様に、設問として不適切と判断された問題についても対象外とした。

データセットの構築にあたっては、第 97 回から第 108 回 JNEP の問題を訓練用データ (Training set)、第 109 回 JNEP の問題をテスト用データ (Test set) として分割した。訓練データには YakugakuQA より各設問に対応する解説文を紐付け、解説が存在しない問題は除外した。最終的に、訓練用 2,421 問、テスト用 165 問のデータセットを構築し (Table 9)、標準化された JSON 形式で整理した (Table 10)。

JNEP は試験の性質上、類似した問題が出題される可能性があり、LLM の性能評価に影響を及ぼすことが懸念される。評価に先立ち、問題文間のテキスト重複度を定量的に評価した。重複度の評価として、二つの観点から実施した。第一に、文字レベルでの重複を評価するため、1-gram から 4-gram までを対象とした頻度分析を実施した。第二に、形態素単位での重複を評価するため、日本語の形態素解析を Janome¹²⁰ を用いて実施した。問題文を形態素に分解した上で、各形態素及び主要な品詞 (名詞、動詞、形容詞) 別の頻度を測定した。

Table 9. Data Overview from JNEP (97th to 109th)

Domain	Number of Questions	
	Training set	Test set
Physics	142	11
Chemistry	39	4
Biology	129	9
Hygiene	320	25
Pharmacology	390	27
Pharmaceutics	300	14
Pathophysiology/Drug Therapy	413	26
Laws/Regulations/Ethics	315	24
Practice	373	25
Total	2,421	165

Table 10. The standardized JSON format for JNEP questions

```
{
  "question_id": "<Unique identifier for the question>",
  "instruction": "<Background information or case vignette applicable to the current question (and potentially others in a sequence)>",
  "input": {
    "current_problem": "<The text of the specific question to be answered>",
    "previous_problems": [
      // List of preceding questions/answers in the same sequence, if applicable
      {
        "problem": "<Text of the previous related question>",
        "answer": "<Correct answer to that previous question>"
      }
    ]
  },
  "choices": [
    "<Text of choice 1>",
    "<Text of choice 2>",
    // ... up to the number of choices for the question
  ],
  "output": "<Identifier(s) corresponding to the correct choice(s)>",
  "explanation": "<Explanatory text detailing the reasoning for the correct answer (primarily used for training data)>",
  "metadata": {
    "examination_session": "<The JNEP administration number (e.g., '109')>",
    "question_number": "<The original question number within the specific JNEP administration>",
    "domain": "<The subject area classification of the question (e.g., 'Pharmacology')>"
  }
}
```

ローカル LLM

4 つのローカル LLM を使用した (Table 11)。Phi-4 (phi-4) は 140 億パラメータという軽量モデルでありながら、高品質な合成データの活用に加え、正答に関係するトークンを特定して最適化する「Pivotal Token Search」を導入することで、高い性能を達成した¹²¹。

残りの 3 つは DeepSeek R1 シリーズである。DeepSeek R1 は推論能力の強化を重視し、DeepSeek-V3-Base を基に強化学習 (Group Relative Policy Optimization) と Supervised Fine-Tuning を組み合わせて訓練されている¹²²。DSR1-Qwen32B (DeepSeek R1 Distill Qwen 32B) は、DeepSeek R1 を Qwen2.5-32B に蒸留したモデルで、英語と中国語のタスクに最適化されている¹²³。

JNEP は日本語で実施されるため、日本語処理に特化したモデルの評価も重要となる。そこで、CyberAgent 社が日本語向けにファインチューニングした 2 つのモデル、CA-DSR1-32B (DeepSeek R1 Distill Qwen 32B fine-tuned for Japanese) と CA-DSR1-14B (DeepSeek R1 Distill Qwen 14B fine-tuned for Japanese) を追加した^{124,125}。

Table 11. Four LLM models used in this study

Short Forms for Models	Model Full Names	Number of Parameters	Developer
phi-4	Phi-4	14 Billion	Microsoft Corporation
DSR1-Qwen32B	DeepSeek R1 Distill Qwen	32 Billion	High-Flyer
CA-DSR1-32B	DeepSeek R1 Distill Qwen fine-tuned for Japanese	32 Billion	High-Flyer and CyberAgent, Inc
CA-DSR1-14B	DeepSeek R1 Distill Qwen fine-tuned for Japanese	14 Billion	High-Flyer and CyberAgent, Inc

ファインチューニング手法

LoRA は、パラメータ効率的なファインチューニング手法である。LoRA は事前学習済みの重み W_0 を固定し、重みの変化量 (ΔW) を表現する低ランク行列のみを学習し、これを元の重みに加算する。

元の重み行列を式 1 のようにした場合 ¹¹⁶、

式 1

$$W_0 \in \mathbb{R}^{d \times k}$$

LoRA は更新量 ΔW に対し低ランク分解を用いて式 2 のように近似する。

式 2

$$\Delta W = BA, \quad \text{where } B \in \mathbb{R}^{d \times r} \text{ and } A \in \mathbb{R}^{r \times k}$$

重要な点は、ランク r が元の次元 d および k よりも大幅に小さいことである（すなわち、 $r \ll \min(d, k)$ ）。ファインチューニングでは、行列 A と B のみが訓練される。その結果、LoRA を適用した層の順伝播は、元の演算 $h = Wx$ (x は入力、 h は出力) を式 3 のように変更する。

式 3

$$h = W_0x + \Delta Wx = W_0x + BAx$$

モデルは低ランクの更新行列 (A および B) のみを学習することで新しいタスクに適用でき、元の重み行列全体をファインチューニングする場合と比較して、訓練可能なパラメータの数を大幅に削減できる。これにより、限られた計算資源でもドメイン特化した適応が可能になり、元のアーキテクチャを変更することなく効率的なファインチューニングを実現する。

前述の 4 つの LLM に対して LoRA を適用した。訓練データを 9:1 の比率でランダムに訓練セットと検証セットに分割した。各モデルは、先行するシーケンスから次のトークンを予測する因果的言語モデリングタスクとしてファインチューニングした。過学習を防ぎ訓練効率を最適化するため、早期停止も追加した。その他の主要な訓練パラメータを Table 12 に示す。

Table 12. Detailed parameters and values used in fine-tuning.

Parameter Category	Parameter Name	Value/Details
LoRA Configuration	Rank (r)	8
	Alpha Scaling Factor (alpha)	16
	Injected Layers	q_proj, k_proj, v_proj, o_proj, gate_proj, up_proj, down_proj
	Dropout Probability	0.1
Training Parameters	Learning Rate	2×10^{-5}
	Maximum Training Duration	20 epochs
	Learning Rate Warmup	The first 10% of the training steps
	Per-device Batch Size	2
	Gradient Accumulation Steps	8
Early Stopping Criteria	Effective Batch Size (with gradient accumulation)	16
	Maximum Input Sequence Length	4096 tokens
	Validation Loss Threshold	1.0
	Plateau Detection	Training terminates if no improvement in validation loss is observed across a predefined number of evaluation steps

プロンプトの設計

全てのモデルで共通のプロンプト構造と入力データを用い、出力形式を統一するプロンプトを先頭に付した。phi-4 と DSR1-Qwen32B には英語で以下の内容を与えた：

The following is a question from the Japanese National Pharmacist Examination. Please ensure your response is returned strictly in the following format: "answer": comma-separated integer(s) (e.g., 1,2,3).

一方、CA-DSR1-14B および CA-DSR1-32B に対しては、同等の内容を日本語で与えた。

次に、プロンプトに問題の主要部分を追加した。問題文として JSON 形式のデータ (Table 10) から指示文 (instruction)、問題文 (current_problem)、選択肢 (choices) を追加した。連続する問題に対しては、前問 (previous_problems) も含めた。ファインチューニング時の訓練データには、正解 (output) と解答根拠 (explanation) をプロンプトに含めた。なお、解答根拠は<think>タグで囲んで示した。

性能評価

LLM の性能評価は、生成された解答に対して 2 段階の解析を行うことで実施した。まず正規表現を用いた自動判定を行い、次に自動判定が困難な箇所を手動で検証した。スコアリングは、正解を 1 点とし、不正解・不完全な出力・解答不明瞭なケースをすべて 0 点とした。主要評価指標には第 109 回 JNEP の正答率を採用した。また、LoRA 適用前後での性能差を統計的に検証するため、マクネマー検定による有意差検定と、ブートストラップ法による信頼区間の推定を行った。なお、科目間の多重比較においてはボンフェローニ補正を適用した。

計算資源

計算は大阪大学の SQUID スーパーコンピュータで実施した。システムは 42 の GPU ノードで構成され、各ノードには Ice Lake アーキテクチャの Intel Xeon Platinum 8368 プロセッサが 2 基搭載されている。プロセッサは基本周波数 2.40 GHz で動作し、1 基あたり 38 物理コアを持つことから、1 ノードあたり合計 76 コアとなる。各 GPU ノードには 512 GB の RAM と 8 基の NVIDIA A100 GPU が搭載されている¹²⁶。評価時のモデル推論には、計算効率と数値安定性のバランスをとるため、半精度浮動小数点 (FP16) を使用した。LoRA 実施時には、モデル性能を維持しながら GPU のメモリ使用量を最適化するため、8 ビット量子化を採用した。

第3節 結果

本研究では、JNEP をベンチマークとして、4つのローカル LLM (phi-4、DSR1-Qwen32B、CA-DSR1-32B、CA-DSR1-14B) を対象に、LoRA 適用前後における正答率の変化を評価した。

まず、JNEP に含まれる類題の重複を評価するため、n-gram 分析および形態素解析を用いて問題間のテキスト重複を分析した。その結果、問題間のテキスト重複は全体として低いことが示された (Figure 10、Figure 11)。文字レベルの n-gram 分析では、1-gram から 4-gram のすべてにおいて出現率の分布がゼロ付近に集中しており、中央値は 1-gram で 0.70%、4-gram で 0.04%と、n-gram の長さに伴い低下する傾向が認められた。形態素解析においても、形態素の出現率の中央値は 0.66%と低く、試験問題間の重複は限定的であることが示唆された。

N-gram Occurrence Ratio Distribution

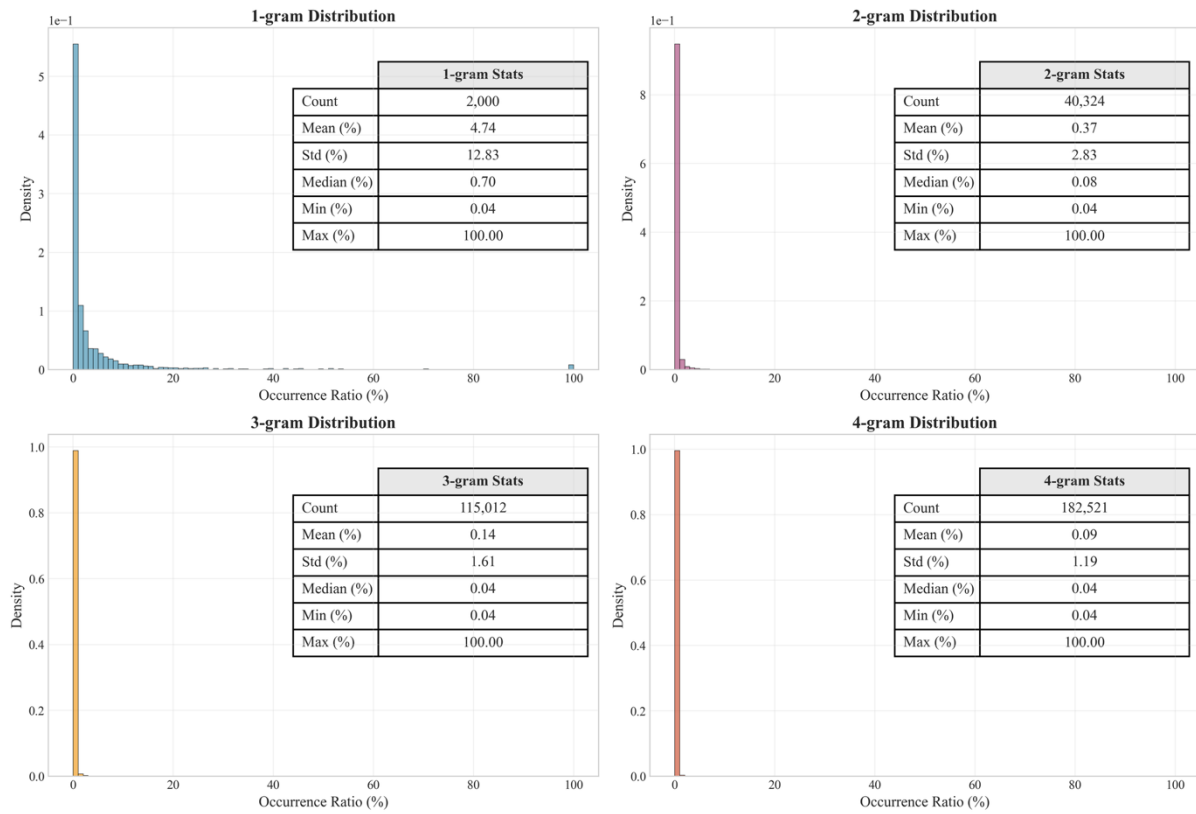


Figure 10. N-gram Occurrence Ratio Distribution

Histograms of occurrence ratios for character n-grams (1-gram to 4-gram) within the JNEP problems. The x-axis denotes the occurrence ratio (%), representing the frequency of each n-gram across all problems, and the y-axis indicates density. All distributions are concentrated near zero, with median ratios of 0.70% (1-gram), 0.08% (2-gram), 0.04% (3-gram), and 0.04% (4-gram), suggesting minimal textual overlap between problems.

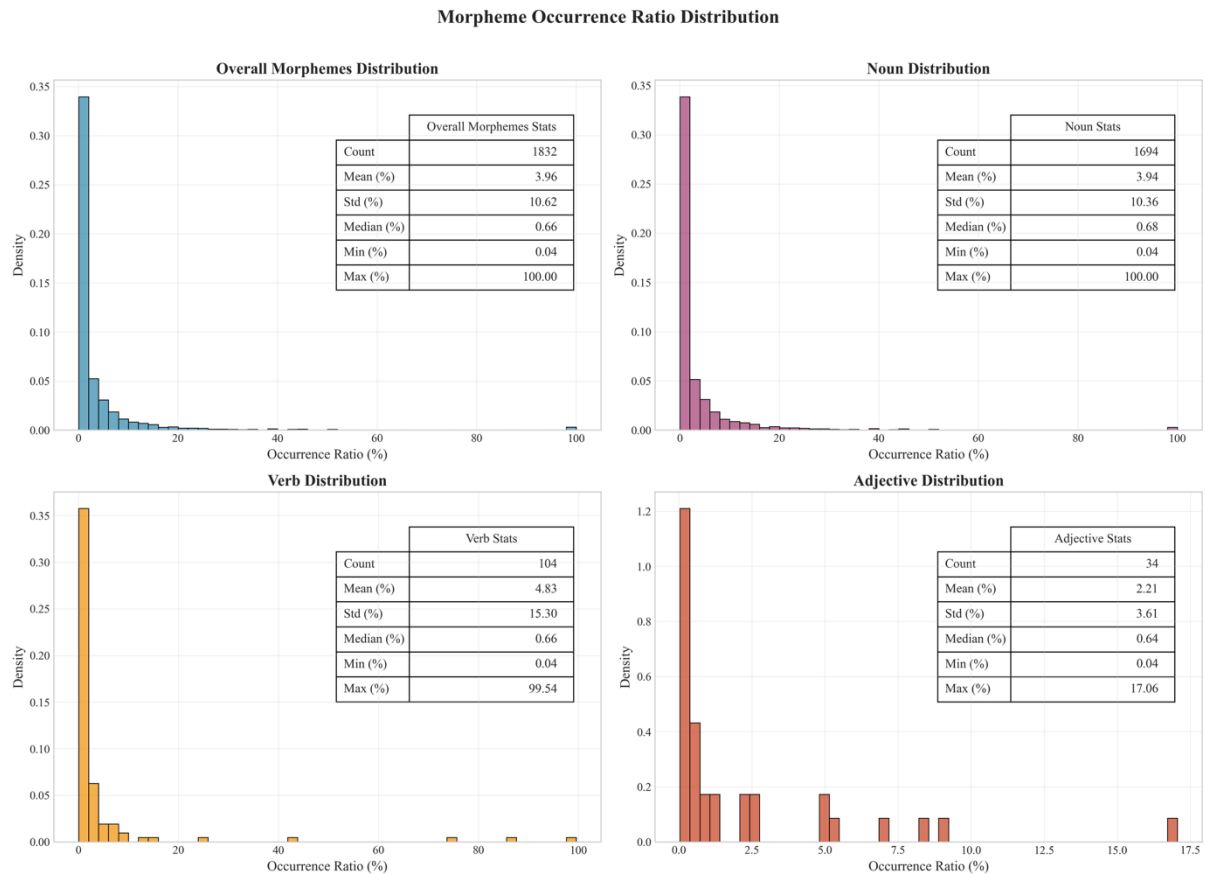


Figure 11. Morpheme Occurrence Ratio Distribution

Histograms showing the distribution of occurrence ratios for morphemes within the JNEP problems, analyzed overall and by part-of-speech (Noun, Verb, Adjective). The x-axis represents the frequency of each morpheme across all problems. All distributions show a strong skew toward zero, with a median ratio of 0.66% for all morphemes, indicating limited structural or semantic duplication.

全体の正答率

次に、4つのローカル LLM について、LoRA 適用前後における全体の正答率を評価した (Table 13)。LoRA 適用前の正答率（ベースライン）は 55.15%から 76.36%の範囲であり、DSR1-Qwen32B が 76.36%と最も高い正答率を示した。LoRA 適用後の正答率は 54.54%から 76.97%の範囲であり、LoRA 適用後も DSR1-Qwen32B が最高値（76.97%）を示した。LoRA による正答率の変化はモデル間で異なり、-0.61%から+5.45%の範囲であった。特に phi-4 では最も大きな改善が認められ、LoRA 適用前後の差は統計的に有意であった（マクネマー検定、 $p = 0.039$ 、95% CI [0.61, 10.30]）。

Table 13. Overall Accuracy Scores for Four Models on the 109th JNEP Test Set

Model	Accuracy (%)		
	Base Model Accuracy, Text-Only No./Total (%)	LoRA Fine-tuned Model Accuracy, Text-Only No./Total (%)	Text-Only Difference (Fine-tuned – Base), (%) [95% CI], p-value
CA-DSR1-14B	91/165 (55.15)	90/165 (54.55)	-0.61 [-7.88, 6.67], $p = 0.873$
CA-DSR1-32B	106/165 (64.24)	112/165 (67.88)	3.64 [-2.42, 9.70], $p = 0.289$
DSR1-Qwen32B	126/165 (76.36)	127/165 (76.97)	0.61 [-5.45, 6.67], $p = 0.841$
phi-4	100/165 (60.61)	109/165 (66.06)	5.45 [0.61, 10.30], $p = 0.039^*$

科目ごとの正答率分析

生物、衛生、病態/薬物治療、実務の科目において、一般に合格水準の目安とされる 60%を超えるスコアが得られた (Table 14)。特に生物では、CA-DSR1-32B と DSR1-Qwen32B が満点を達成し、LoRA 適用後もその成績を維持した。

LoRA の効果を評価するため、マクネマー検定を行い、多重比較に対してはボンフェローニ補正を適用した。これらの厳格な基準では統計的有意性は認められず、ブートストラップ信頼区間も 0 を含むことが多かったが、複数の科目で意味のある改善が観察された。

衛生学領域では、全てのモデルで LoRA 適用後に一貫した正答率向上が見られ、その増加幅は 4.0%から 8.0%であった。

薬理、薬剤、病態/薬物治療といった薬学系の科目では、LoRA 後に一部のモデルで改善が認められた。phi-4 では薬理 (+11.1%、95%信頼区間[-3.7, 18.5]) および薬剤 (+14.3%、95%信頼区間[14.3, 14.3]) で正答率向上が見られ、CA-DSR1-32B も病態/薬物治療で+11.5% (95%信頼区間[-3.8, 26.9]) の増加を示した。

法規/制度/倫理では、日本語に公式対応していない DSR1-Qwen32B が比較的良好なベースライン正答率 (54%) を示し、LoRA 適用後には+16.7% (95%信頼区間[16.7, 16.7]) と正答率が向上した。

Table 14. Domain-Specific Accuracy and Changes Attributable to LoRA Fine-Tuning on the Test Set (the 109th JNEP)

Domain	n	CA-DSR1-14B			CA-DSR1-32B			DSR1-Qwen32B			phi-4		
		Base (n)	Fine-tuned (n)	$\Delta(\%)$ [95% CI]	Base (n)	Fine-tuned (n)	$\Delta(\%)$ [95% CI]	Base (n)	Fine-tuned (n)	$\Delta(\%)$ [95% CI]	Base (n)	Fine-tuned (n)	$\Delta(\%)$ [95% CI]
Physics	11	7	6	-9.1 [-27.3, 27.3]	9	8	-9.1 [-9.1, -9.1]	8	9	9.1 [9.1, 9.1]	8	7	-9.1 [-9.1, -9.1]
Chemistry	4	2	0	-50.0	1	3	50.0 [50.0, 50.0]	4	2	-50.0	3	3	0.0
Biology	9	7	9	22.2	9	9	0.0	9	9	0.0	6	6	0.0
Hygiene	25	15	16	4.0 [-20.0, 28.0]	18	19	4.0 [-12.0, 20.0]	20	22	8.0 [-16.0, 24.0]	15	17	8.0 [-8.0, 16.0]
Pharmacology	27	10	8	-7.4 [-22.2, 7.4]	17	17	0.0 [-14.8, 14.8]	23	20	-11.1 [-11.1, 11.1]	11	14	11.1 [-3.7, 18.5]
Pharmaceutics	14	7	7	0.0 [-14.3, 14.3]	7	8	7.1 [-21.4, 21.4]	9	8	-7.1 [-7.1, -7.1]	7	9	14.3 [14.3, 14.3]
Pathophysiology/ Drug Therapy	26	16	15	-3.8 [-19.2, 11.5]	18	21	11.5 [-3.8, 26.9]	24	22	-7.7 [-7.7, -7.7]	19	20	3.8 [3.8, 3.8]
Laws/Regulations/ Ethics	24	12	13	4.2 [-12.5, 12.5]	12	12	0.0 [-16.7, 16.7]	13	17	16.7 [16.7, 16.7]	14	17	12.5 [-4.2, 20.8]
Practice	25	15	16	4.0 [-12.0, 20.0]	15	15	0.0 [-16.0, 16.0]	16	18	8.0 [-8.0, 24.0]	17	16	-4.0 [-4.0, -4.0]

第4節 考察

ローカル LLM の性能と、LoRA を用いたファインチューニングの有効性を JNEP で探索的に評価した。評価したローカル LLM のうち、CA-DSR1-32B と DSR1-Qwen32B は、ベースライン正答率で第 109 回 JNEP の合格基準（60.87%）を上回った。さらに、LoRA 適用により phi-4 モデルも合格基準を超え、さらに、パラメータ数の大きな CA-DSR1-32B のベースライン正答率を上回った。phi-4 は約 140 億という比較的小規模なパラメータ数にもかかわらず一定水準の正答率を示した。この結果は、リソース制約のある環境においても、ローカル LLM が一定の性能を発揮し得る可能性を示唆するものである。また、英語および中国語で学習された DSR1-Qwen32B が日本語試験において高い性能を示した点は、多言語 LLM の言語横断的能力を報告した先行研究の知見と一致する¹²⁷。

本研究ではモデルのパラメータ数と JNEP の正答率には正の関連が見られ、320 億パラメータの LLM が 140 億パラメータの LLM を上回った。この知見は LLM におけるスケーリング則と一致しており、より大きなパラメータ数が JNEP のような多面的な知識タスクにおいて、知識獲得・統合能力を向上させる可能性を示唆している^{128,129}。ただし、近年の研究ではモデル圧縮などの技術により、小規模モデルが大規模モデルと同等の性能を達成する可能性も示されており¹³⁰、パラメータ数が決定的要因ではないことにも留意が必要である。

LoRA は、全体としてモデル正答率を向上させたが、効果はモデルによって一様ではなかった。最も顕著な改善が見られたのは phi-4 であり、LoRA 適用後には合格基準を超え、CA-DSR1-32B のベースライン正答率を上回った。一方、DSR1-Qwen32B と CA-DSR1-14B では性能の変化はわずかであった。ファインチューニングの効果がモデルやタスク間で異なるという結果は、先行研究¹³¹の知見とも一致する。これらの結果は、LoRA が薬学領域において実行可能で効果的な戦略である可能性を示す一方で、ファインチューニング手法についてさらなる検討が必要である。

科目別に見ると、生物、衛生、病態/薬物治療、実務の各科目において、全モデルが合格水準の目安となる 60%を超えるベースライン正答率を達成した。これらの科目で求められる知識は、一般的な理系・医学分野の知識と重なる部分が多く、LLM が事前学習で獲得した知識を活用しやすいため、高い性能に寄与したと考えられる。特に衛生では、全てのモデルで LoRA 適用後に正答率向上が見られた。衛生は公衆衛生、環境科学、疫学など幅広いトピックの知識問題を含むため、LoRA による追加学習で関連知識が効果的に補強されやすかったと推測される。

薬学系の科目である薬理・薬剤、および病態/薬物治療において、一部のモデルは LoRA 適用後に正答率向上を示した。特に phi-4 では薬理 (+11.1%) および薬剤 (+14.3%) で改善が見られた。また、CA-DSR1-32B では病態/薬物治療 (+11.5%) で改善が認められた。これらの結果は、ファインチューニングが専門領域での性能を向上させるという先行研究の知見と一致しており、LoRA が薬学の専門知識の補強に有効である可能性を示唆している^{129,132-135}。

法規・倫理の科目では、日本語に正式対応していない DSR1-Qwen32B でも比較的良好なベースライン正答率 (54%) を示し、LoRA 適用後には+16.7%と大幅な改善を達成した。事前学習の段階で日本の法規制に関する一定の知識を有しており、LoRA により知識が補強された可能性を示唆する。これは、各国の法律データで LLM をファインチューニングすることで関連タスクの精度が向上するという先行研究の知見と一致する^{107,136}。

本章で報告した研究には、いくつかの限界が存在する。第一に、本研究では画像問題を扱うことができないため、評価対象をテキスト問題に限定した。第二に、LLM の事前学習データに関する透明性の課題が挙げられる。事前学習データの詳細は公開されていないため、評価に用いた試験問題が事前学習に含まれていなかったかを厳密に確認することは困難である。特に知識カットオフ日の設定を考慮すると、学習データへの混入の可能性は完全には排除できず、本評価を解釈する上での留意点の一つとなる。第三に、ローカル LLM はセキュリティ面で一定の利点を有する一方で、性能面ではクラウド型 LLM に劣る傾向がある。近年、クラウド型サービスも多様化していることから、ローカル LLM の相対的な位置づけや優位性については、今後も継続的な検討が必要である。第四に、本研究は定義された実験条件下で実施した探索的研究であり、評価対象を 4 つのモデルに限定している。そのため、本研究の結果は当該条件およびモデルの範囲内で解釈されるべきであり、他のモデルへの一般化可能性や異なる条件下での性能については未検証である。第五に、本研究では JNEP の正答率のみを評価指標としており、LLM の安全性や倫理的側面、ならびに医療への応用可能性に関する検討は含まれていない。これらの限界を踏まえ、今後はデータリークの検証や、クラウド型を含む多様な LLM を対象とした比較、さらに安全性・倫理性の観点を含めた包括的な評価が求められる。

第5節 小括

第2章で開発した AEST において、シグナル検出結果を臨床応用へと発展させるための情報提供機能を追加することは、臨床現場での活用促進につながると期待される。特に、シグナルの臨床的意義を医学的・薬学的観点から提示する機能は、ユーザーの安全性研究に対する理解を深め、実臨床での意思決定支援に寄与する。

本章では、情報提供機能の実装に向けた予備的かつ探索的検討として、LLM、特にセキュアな環境でも運用可能なローカル LLM の薬学領域における性能について評価を行った。また、薬学に特化した LLM 構築に向けて、ファインチューニング手法に関する基礎的な検討を行った。JNEP を用いた性能評価の結果、合格基準を上回るローカル LLM が存在したほか、ファインチューニングによって新たに合格基準を超える LLM も確認された。これらの結果は、ローカル LLM が医学・薬学的観点からの情報提供を目的とした応用において、一定の可能性を有することを示唆する。

AEST の情報提供機能の開発において、薬学的知識を備えたローカル LLM は有効な情報提供基盤となり得る (Figure 12)。ローカル LLM を AEST に組み込むことで、ROR などの統計指標の解釈上の留意点や臨床上の注意事項を文章で提示するとともに、関連する添付文書情報や類似症例報告を自動的に整理・提示し、シグナル解釈を支援する機能の実装が期待される。これにより、利用者のデータ理解が促進され、有害事象情報の利活用につながると考えられる。今後は、本章で得られた成果を基盤として、AEST の情報提供機能を、臨床現場における意思決定支援に寄与する形へと段階的に発展させていく予定である。

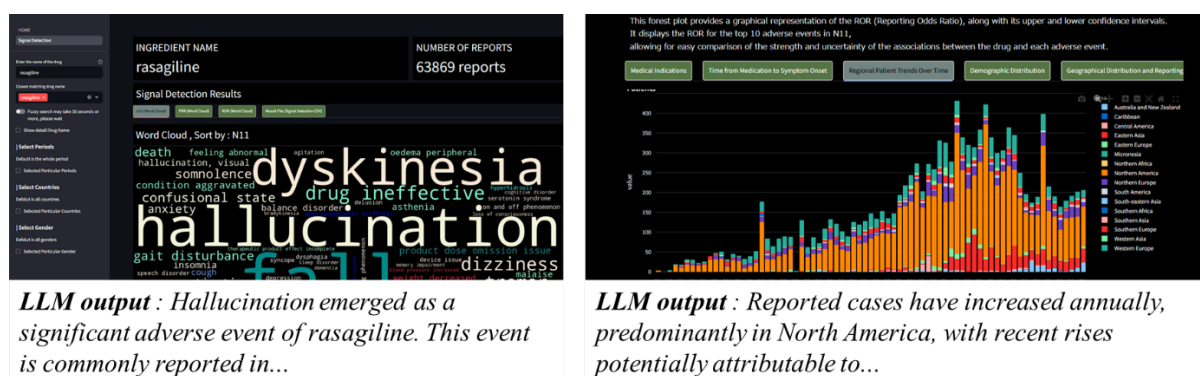


Figure 12. Conceptual Framework for Integrating AEST Visualizations with LLM-Generated Interpretations.

結論

本論文では、主要な自発報告データベースの1つである FAERS を活用した医薬品安全性評価の推進を目指し、3つのアプローチから研究を実施した。

第1章では、同一薬理作用機序を有する MAO-B 阻害薬（セレギリン、ラサギリン、サフィナミド）を対象に、FAERS を用いた安全性比較を実施した。比較解析の結果、薬理作用が類似する医薬品間においても、化学構造や代謝経路などの差異が ADR の発現に影響を及ぼす可能性が示された。特に、セレギリンでの肝胆道系障害、ラサギリンでの血管障害や腫瘍関連シグナル、サフィナミドでの眼障害といった、医薬品ごとに特異的なシグナルが検出された。MAO-B 阻害薬における安全性比較の結果は、パーキンソン病治療における治療薬選択の客観的根拠を補強するものである。また、本手法は他の薬効群にも応用可能であり、医薬品安全性評価における新たな比較解析アプローチとして有用性を示した。

第2章では、FAERS の品質および解析環境に関する課題に対応するため、Web アプリケーション「AEST (Adverse Events Signal-detection Tool)」を開発した。AEST は、標準化された FAERS データを基盤とし、5,721 種類の医薬品に対して高速かつ再現性の高いシグナル検出を実現する。加えて、直感的なインターフェースを備えており、簡易な操作でシグナル検出を実行できる環境を提供する。AEST の開発により、医薬品安全性研究の裾野が拡大し、多角的な検討の促進が期待される。

第3章では、AEST の解析結果の臨床的価値を高め、その活用を支援する機能拡張の一環として、ローカル LLM 技術の導入可能性について探索的に検討した。4つのローカル LLM を対象に、JNEP を指標として薬学領域における性能を評価するとともに、LoRA によるファインチューニングが性能変化に及ぼし得る影響について基礎的な検討を行った。その結果、限られた計算資源下においても、薬学領域に特化したモデルを構築できる可能性が示唆された。本検討は、ローカル LLM を用いることで、セキュリティを確保しつつ専門性の高い情報提供機能を AEST に統合し得る可能性を示すものであり、AEST における高度な解析支援機能の実装に向けた基盤となり得る。

本研究の3つの成果は、独立した知見を提供するとともに、FAERS の利活用を基盤とした医薬品安全性研究の発展に寄与するものである。第1章で確立した医薬品比較解析手法は、第2章の AEST を通じて広く活用可能となり、第3章で検証したローカル LLM は AEST の活用支援機能として統合することで、より高度な安全性評価環境の構築に向けた可能性を示した。本研究で得られた知見と開発成果は、ファーマコビジランス研究の推進を通じて、国民の健康増進および医療の質の向上に資すると考える。

謝辞

本研究を学位論文としてまとめるにあたり、終始ご指導とご鞭撻を賜りました大阪大学大学院薬学研究科医療薬学分野 教授 池田 賢二先生に、心より深く感謝申し上げます。

また、本研究において多大なるご指導とご助言を賜りました同研究科 特任教授 上田 幹子先生、准教授 大石 美奈子先生、助教 幡生 あすか先生に、厚く御礼申し上げます。

さらに、研究の遂行に際し、温かいご指導とご協力を賜りました大阪大学大学院薬学研究科量子生命情報薬学分野 教授 福澤 薫先生、講師 高谷 大輔先生、助教 田 雨時先生に、深く感謝いたします。

あわせて、本研究の推進にあたり多方面にわたりご助力をいただきました大阪大学大学院薬学研究科 名誉教授 高木 達也先生に、心より感謝申し上げます。

最後に、日々の研究活動を通じてご協力と有益な議論をいただいた大阪大学大学院薬学研究科医療薬学分野および量子生命情報薬学分野の皆様、厚く御礼申し上げます。

参考文献

1. Kommu, S., Carter, C. & Whitfield, P. Adverse Drug Reactions. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2025).
2. Cosgrave, N. *et al.* Hospital admissions due to adverse drug reactions and adverse drug events in older adults: a systematic review. *Age Ageing* 54, afaf231 (2025).
3. Kauppila, M., Backman, J. T., Niemi, M. & Lapatto-Reiniluoto, O. Drug-related deaths in a university hospital: Comparison to previous decades. *Basic & Clinical Pharmacology & Toxicology* 134, 165-174 (2024).
4. Laurens M. B. *et al.* Definition and assessment of adverse events. *Trials* <https://bio-protocol.org/exchange/minidetail?id=827169&type=30> (2016).
5. Sahu, R. K., Yadav, R., Prasad, P., Roy, A. & Chandrakar, S. Adverse drug reactions monitoring: prospects and impending challenges for pharmacovigilance. *Springerplus* 3, 695 (2014).
6. Caparrotta, T. M., Dear, J. W., Colhoun, H. M. & Webb, D. J. Pharmacoepidemiology: Using randomised control trials and observational studies in clinical decision - making. *Br J Clin Pharmacol* 85, 1907-1924 (2019).
7. Norén, G. N. & Edwards, I. R. Modern methods of pharmacovigilance: detecting adverse effects of drugs. *Clin Med (Lond)* 9, 486-489 (2009).
8. Potter, E., Reyes, M., Naples, J. & Dal Pan, G. FDA Adverse Event Reporting System (FAERS) Essentials: A Guide to Understanding, Applying, and Interpreting Adverse Event Data Reported to FAERS. *Clinical Pharmacology & Therapeutics* 118, 567-582 (2025).
9. Research, C. for D. E. and. FDA Adverse Event Reporting System (FAERS) Public Dashboard. *FDA* <https://www.fda.gov/drugs/fdas-adverse-event-reporting-system-faers/fda-adverse-event-reporting-system-faers-public-dashboard> (2023).
10. Alomar, M., Tawfiq, A. M., Hassan, N. & Palaian, S. Post marketing surveillance of suspected adverse drug reactions through spontaneous reporting: current status, challenges and the future. *Therapeutic Advances in Drug Safety* 11, 2042098620938595 (2020).
11. Masiliūnienė, G., Stankevičius, E. & Kaduševičius, E. Trends in adverse drug reaction reporting in eight selected countries after the implementation of new pharmacovigilance regulation in 2012: a joinpoint regression analysis. *Eur J Clin Pharmacol* 81, 1287-1299 (2025).
12. Al Dweik, R., Stacey, D., Kohen, D. & Yaya, S. Factors affecting patient reporting of adverse drug reactions: a systematic review. *British Journal of Clinical Pharmacology* 83, 875-883 (2017).
13. Blenkinsopp, A., Wilkie, P., Wang, M. & Routledge, P. A. Patient reporting of suspected adverse drug reactions: a review of published literature and international experience. *British Journal of Clinical Pharmacology* 63, 148-156 (2007).
14. Jiao, X.-F. *et al.* Adverse drug reaction signal detection methods in spontaneous reporting system: A systematic review. *Pharmacoepidemiology and Drug Safety* 33, e5768 (2024).
15. MOKBEL, K., DANIELS, R., WEEDON, M. N. & JACKSON, L. A Comparative Safety Analysis of Medicines Based on the UK Pharmacovigilance and General Practice Prescribing Data in England. *In*

- Vivo* 36, 780-800 (2022).
16. Maciejewski, M. *et al.* Reverse translation of adverse event reports paves the way for de-risking preclinical off-targets. *eLife* 6, e25818 (2017).
 17. Soldatos, T. G., Kim, S., Schmidt, S., Lesko, L. J. & Jackson, D. B. Advancing drug safety science by integrating molecular knowledge with post-marketing adverse event reports. *CPT: Pharmacometrics & Systems Pharmacology* 11, 540-555 (2022).
 18. Nango, D., Sekizuka, T., Goto, M. & Echizen, H. Analysis of Information on Drug Adverse Reactions Using U.S. Food and Drug Administration Adverse Event Reporting System (FAERS). *Yakugaku Zasshi* 142, 341-344 (2022).
 19. Rocca, E., Copeland, S. & Ralph Edwards, I. Pharmacovigilance as Scientific Discovery: An Argument for Trans-Disciplinarity. *Drug Saf* 42, 1115-1124 (2019).
 20. Yamamoto, M. The Role of a Pharmacist and Undergraduate Pharmacy Education with Special Reference to the Adverse Drug Reaction Reporting System in Japan. *ResearchGate*
https://www.researchgate.net/publication/314937630_The_Role_of_a_Pharmacist_and_Undergraduate_Pharmacy_Education_with_Special_Reference_to_the_Adverse_Drug_Reaction_Reporting_System_in_Japan.
 21. Fan, M. *et al.* Comprehensive safety evaluation of a novel multitargeting compound XYY-CP1106: A candidate for Alzheimer's disease. *Biomedicine & Pharmacotherapy* 176, 116786 (2024).
 22. Lee, H. Y. *et al.* Astrocytic monoamine oxidase B (MAOB)-gamma-aminobutyric acid (GABA) axis as a molecular brake on repair following spinal cord injury. *Sig Transduct Target Ther* 10, 295 (2025).
 23. Ou, Z. *et al.* Global Trends in the Incidence, Prevalence, and Years Lived With Disability of Parkinson's Disease in 204 Countries/Territories From 1990 to 2019. *Front Public Health* 9, 776847 (2021).
 24. Armstrong, M. J. & Okun, M. S. Diagnosis and Treatment of Parkinson Disease: A Review. *JAMA* 323, 548-560 (2020).
 25. Finberg, J. P. M. & Rabey, J. M. Inhibitors of MAO-A and MAO-B in Psychiatry and Neurology. *Frontiers in Pharmacology* 7, 340 (2016).
 26. Tábi, T., Vécsei, L., Youdim, M. B., Riederer, P. & Szökö, É. Selegiline: a molecule with innovative potential. *J Neural Transm* 127, 831-842 (2020).
 27. Moore, J. J. & Saadabadi, A. *Selegiline*. *StatPearls [Internet]* (StatPearls Publishing, 2021).
 28. Fozard, J. R., Zreika, M., Robin, M. & Palfreyman, M. G. The functional consequences of inhibition of monoamine oxidase type B: comparison of the pharmacological properties of L-deprenyl and MDL 72145. *Naunyn Schmiedebergs Arch Pharmacol* 331, 186-193 (1985).
 29. Okuda, C., Segal, D. S. & Kuczenski, R. Deprenyl alters behavior and caudate dopamine through an amphetamine-like action. *Pharmacol Biochem Behav* 43, 1075-1080 (1992).
 30. Engberg, G., Elebring, T. & Nissbrandt, H. Deprenyl (selegiline), a selective MAO-B inhibitor with active metabolites; effects on locomotor activity, dopaminergic neurotransmission and firing rate of nigral dopamine neurons. *J Pharmacol Exp Ther* 259, 841-847 (1991).

31. Churchyard, A., Mathias, C. J., Boonkongchuen, P. & Lees, A. J. Autonomic effects of selegiline: possible cardiovascular toxicity in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 63, 228-234 (1997).
32. Stryjer, R., Klein, C., Treves, T. A. & Rabey, J. M. The effects of acute loading with levodopa and levodopa with selegiline on blood pressure and plasma norepinephrine levels in chronic Parkinson's disease patients. *Acta Neurologica Scandinavica* 111, 89-94 (2005).
33. Montastruc, J. L. *et al.* Adverse drug reactions to selegiline: a review of the French pharmacovigilance database. *Clin Neuropharmacol* 23, 271-275 (2000).
34. Yasar, S. *et al.* A comparison of drug-seeking behavior maintained by D-amphetamine, L-deprenyl (selegiline), and D-deprenyl under a second-order schedule in squirrel monkeys. *Psychopharmacology (Berl)* 183, 413-421 (2006).
35. Yasar, S. *et al.* Metabolic transformation plays a primary role in the psychostimulant-like discriminative-stimulus effects of selegiline [(R)-(-)-deprenyl]. *J Pharmacol Exp Ther* 317, 387-394 (2006).
36. Youdim, M. B. H., Gross, A. & Finberg, J. P. M. Rasagiline [N-propargyl-1R(+)-aminoindan], a selective and potent inhibitor of mitochondrial monoamine oxidase B. *British Journal of Pharmacology* 132, 500-506 (2001).
37. Müller, T., Hoffmann, J. A., Dimpfel, W. & Oehlwein, C. Switch from selegiline to rasagiline is beneficial in patients with Parkinson's disease. *J Neural Transm* 120, 761-765 (2013).
38. Goren, T., Adar, L., Sasson, N. & Weiss, Y. M. Clinical pharmacology tyramine challenge study to determine the selectivity of the monoamine oxidase type B (MAO-B) inhibitor rasagiline. *J Clin Pharmacol* 50, 1420-1428 (2010).
39. DailyMed - AZILECT- rasagiline mesylate tablet.
<https://dailymed.nlm.nih.gov/dailymed/lookup.cfm?setid=a40d0e73-3f9f-4b01-979d-402c9cdaeb60>.
40. Kulisevsky, J. Emerging role of safinamide in Parkinson's disease therapy. *European Neurological Review* 9, 108-112 (2014).
41. deSouza, R. M. & Schapira, A. Safinamide for the treatment of Parkinson's disease. *Expert Opinion on Pharmacotherapy* 18, 937-943 (2017).
42. FAERS Quarterly Data Extract Files. <https://fis.fda.gov/extensions/FPD-QDE-FAERS/FPD-QDE-FAERS.html>.
43. ISO - ISO 3166 – Country Codes. *ISO* <https://www.iso.org/iso-3166-country-codes.html>.
44. KEGG DRUG Database. <https://www.genome.jp/kegg/drug/>.
45. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28, 27-30 (2000).
46. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci* 28, 1947-1951 (2019).
47. Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M. & Tanabe, M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res* 49, D545-D551 (2021).

48. MedDRA. <https://www.meddra.org/>.
49. UNSD – Methodology. <https://unstats.un.org/unsd/methodology/m49/>.
50. Rothman, K. J., Lanes, S. & Sacks, S. T. The reporting odds ratio and its advantages over the proportional reporting ratio. *Pharmacoepidemiol Drug Saf* 13, 519-523 (2004).
51. Fukazawa, C., Hinomura, Y., Kaneko, M. & Narukawa, M. Significance of data mining in routine signal detection: Analysis based on the safety signals identified by the FDA. *Pharmacoepidemiology and Drug Safety* 27, 1402-1408 (2018).
52. Batagelj, V. & Bren, M. Comparing resemblance measures. *Journal of Classification* 12, 73-90 (1995).
53. Iranzo, A., Santamaria, J. & Tolosa, E. The clinical and pathophysiological relevance of REM sleep behavior disorder in neurodegenerative diseases. *Sleep Medicine Reviews* 13, 385-401 (2009).
54. Plastino, M. *et al.* Effects of safinamide on REM sleep behavior disorder in Parkinson disease: A randomized, longitudinal, cross-over pilot study. *Journal of Clinical Neuroscience* 91, 306-312 (2021).
55. DailyMed - AMPHETAMINE SULFATE- amphetamine tablet.
<https://dailymed.nlm.nih.gov/dailymed/drugInfo.cfm?setid=828c84a6-a317-4e4d-9521-0c3865ff732f>.
56. Van Den Eeden, S. K. *et al.* Incidence of Parkinson’ s disease: variation by age, gender, and race/ethnicity. *Am J Epidemiol* 157, 1015-1022 (2003).
57. McLean, G., Hindle, J. V., Guthrie, B. & Mercer, S. W. Co-morbidity and polypharmacy in Parkinson’ s disease: insights from a large Scottish primary care database. *BMC Neurol* 17, 126 (2017).
58. Fowler, J. S. *et al.* Evidence that Formulations of the Selective MAO-B Inhibitor, Selegiline, which Bypass First-Pass Metabolism, also Inhibit MAO-A in the Human Brain. *Neuropsychopharmacology* 40, 650-657 (2015).
59. Kyritsi, K. *et al.* Modulation of the Tryptophan Hydroxylase 1/Monoamine Oxidase-A/5-Hydroxytryptamine/5-Hydroxytryptamine Receptor 2A/2B/2C Axis Regulates Biliary Proliferation and Liver Fibrosis During Cholestasis. *Hepatology* 71, 990-1008 (2020).
60. Mao, B. *et al.* The janus face of serotonin: Regenerative promoter and chronic liver disease aggravator. *Heliyon* 10, e30703 (2024).
61. Hidestrand, M. *et al.* CYP2B6 and CYP2C19 as the major enzymes responsible for the metabolism of selegiline, a drug used in the treatment of Parkinson’ s disease, as revealed from experiments with recombinant enzymes. *Drug Metab Dispos* 29, 1480-1484 (2001).
62. Jost, W. H. A critical appraisal of MAO-B inhibitors in the treatment of Parkinson’ s disease. *J Neural Transm (Vienna)* 129, 723-736 (2022).
63. Kubota, T. , Chiba, K. & Iga, T. Frequency distribution of CYP2C19, CYP2D6, and CYP2C9 mutant-alleles in several different populations. *Drug Metabolism and Pharmacokinetics* 16, 69-74 (2001).
64. Mizuno, Y. *et al.* A Randomized Double-Blind Placebo-Controlled Phase III Trial of Selegiline Monotherapy for Early Parkinson Disease. *Clin Neuropharmacol* 40, 201-207 (2017).

65. Parkinson Study Group. A Controlled Trial of Rasagiline in Early Parkinson Disease: The TEMPO Study. *Archives of Neurology* 59, 1937-1943 (2002).
66. Stern, M. B. *et al.* Double-blind, randomized, controlled trial of rasagiline as monotherapy in early Parkinson's disease patients. *Movement Disorders* 19, 916-923 (2004).
67. Parkinson Study Group. A Randomized Placebo-Controlled Trial of Rasagiline in Levodopa-Treated Patients With Parkinson Disease and Motor Fluctuations: The PRESTO Study. *Archives of Neurology* 62, 241-248 (2005).
68. Rascol, O. *et al.* Rasagiline as an adjunct to levodopa in patients with Parkinson's disease and motor fluctuations (LARGO, Lasting effect in Adjunct therapy with Rasagiline Given Once daily, study): a randomised, double-blind, parallel-group trial. *The Lancet* 365, 947-954 (2005).
69. Parkinson Study Group. A Controlled, Randomized, Delayed-Start Study of Rasagiline in Early Parkinson Disease. *Archives of Neurology* 61, 561-566 (2004).
70. Olanow, C. W. *et al.* A Double-Blind, Delayed-Start Trial of Rasagiline in Parkinson's Disease. *N Engl J Med* 361, 1268-1278 (2009).
71. Zhang, X., Guarin, D., Mohammadzadehhonarvar, N., Chen, X. & Gao, X. Parkinson's disease and cancer: a systematic review and meta-analysis of over 17 million participants. *BMJ Open* 11, e046329 (2021).
72. Johannes, C. B. *et al.* The risk of melanoma with rasagiline compared with other antiparkinsonian medications: A retrospective cohort study in the United States medicare database. *Pharmacoepidemiol Drug Saf* 31, 643-651 (2022).
73. Wasan, H., Singh, D. & Kh, R. Safinamide in neurological disorders and beyond: Evidence from preclinical and clinical studies. *Brain Research Bulletin* 168, 165-177 (2021).
74. deMarcaida, J. A. *et al.* Effects of tyramine administration in Parkinson's disease patients treated with selective MAO-B inhibitor rasagiline. *Movement Disorders* 21, 1716-1721 (2006).
75. Luppi, P.-H., Clément, O., Valencia Garcia, S., Brischoux, F. & Fort, P. New aspects in the pathophysiology of rapid eye movement sleep behavior disorder: the potential role of glutamate, gamma-aminobutyric acid, and glycine. *Sleep Med* 14, 714-718 (2013).
76. Ronconi, G. *et al.* Prescription Pattern of Monoamine Oxidase B Inhibitors Combined with Levodopa: A Retrospective Observational Analysis of Italian Healthcare Administrative Databases. *Drugs - Real World Outcomes* 9, 391-401 (2022).
77. Berman, S. M., Kuczenski, R., McCracken, J. T. & London, E. D. Potential Adverse Effects of Amphetamine Treatment on Brain and Behavior: A Review. *Mol Psychiatry* 14, 123-142 (2009).
78. Sinha, A., Lewis, O., Kumar, R., Yeruva, S. L. H. & Curry, B. H. Adult ADHD Medications and Their Cardiovascular Implications. *Case Rep Cardiol* 2016, 2343691 (2016).
79. Derlet, R. W., Rice, P., Horowitz, B. Z. & Lord, R. V. Amphetamine toxicity: experience with 127 cases. *J Emerg Med* 7, 157-161 (1989).
80. Kalant, H. & Kalant, O. J. Death in amphetamine users: causes and rates. *Can Med Assoc J* 112, 299-304 (1975).

81. Coyle, J. T. & Snyder, S. H. Antiparkinsonian drugs: inhibition of dopamine uptake in the corpus striatum as a possible mechanism of action. *Science* 166, 899-901 (1969).
82. Hazell, L. & Shakir, S. A. W. Under-reporting of adverse drug reactions : a systematic review. *Drug Saf* 29, 385-396 (2006).
83. Khouri, C. *et al.* A meta-epidemiological study found lack of transparency and poor reporting of disproportionality analyses for signal detection in pharmacovigilance databases. *Journal of Clinical Epidemiology* 139, 191-198 (2021).
84. Ge, C., Jin, L., Tian, J.-J., Yang, N. & Xu, J. Comparative safety analysis of lacosamide and perampanel in epilepsy management: insights from FAERS database. *Front. Pharmacol.* 15, (2024).
85. Mu, L., Xu, J., Ye, X., Jiang, Y. & Yi, Z. Comparative safety signals of dopamine agonists: psychiatric and cardiovascular risks derived from FDA adverse event reporting system (FAERS) data. *BMC Pharmacol Toxicol* 26, 54 (2025).
86. AERS spider: an online interactive tool to mine statistical associations in Adverse Event Reporting System - Grigoriev - 2014 - Pharmacoepidemiology and Drug Safety - Wiley Online Library. <https://onlinelibrary.wiley.com/doi/10.1002/pds.3561>.
87. AERSMine: Visualize Differential Therapeutic Outcomes. <https://research.cchmc.org/aers/home>.
88. OpenVigil Pharmacovigilance Search Engines. <https://openvigil.sourceforge.net/#>.
89. Search FDA FAERS and MAUDE databases (drug and medical-device adverse events). <https://www.fdatable.com/>.
90. Lenart, A., Aburto, J. M., Stockmarr, A. & Vaupel, J. W. The human longevity record may hold for decades. Preprint at <https://doi.org/10.48550/arXiv.1809.03732> (2018).
91. Fusaroli, M. *et al.* Enhancing Transparency in Defining Studied Drugs: The Open-Source Living DiAna Dictionary for Standardizing Drug Names in the FAERS. *Drug Saf* 47, 271-284 (2024).
92. OSF. <https://osf.io/zqu89/files>.
93. Folk, M., Heber, G., Koziol, Q., Pourmal, E. & Robinson, D. An overview of the HDF5 technology suite and its applications. in *Proceedings of the EDBT/ICDT 2011 Workshop on Array Databases* 36-47 (Association for Computing Machinery, New York, NY, USA, 2011). doi:10.1145/1966895.1966900.
94. Virtual Datasets (VDS) – h5py 3.15.1 documentation. <https://docs.h5py.org/en/stable/vds.html>.
95. Streamlit · A faster way to build and share data apps. <https://streamlit.io/> (2021).
96. Moritz, P. *et al.* Ray: A Distributed Framework for Emerging AI Applications. Preprint at <https://doi.org/10.48550/arXiv.1712.05889> (2018).
97. Evans, S. J. W., Waller, P. C. & Davis, S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiology and Drug Safety* 10, 483-486 (2001).
98. Senger, C., Kaltschmidt, J., Schmitt, S. P. W., Pruszydlo, M. G. & Haefeli, W. E. Misspellings in drug information system queries: characteristics of drug name spelling errors and strategies for their prevention. *Int J Med Inform* 79, 832-839 (2010).

99. Waruru, A. *et al.* Where No Universal Health Care Identifier Exists: Comparison and Determination of the Utility of Score-Based Persons Matching Algorithms Using Demographic Data. *JMIR Public Health and Surveillance* 4, e10436 (2018).
100. Ma, E. NLP Augmentation. (2019).
101. Search FDA FAERS and MAUDE databases (drug and medical-device adverse events).
<https://www.fda.gov/oc/maude/>.
102. Grigoriev, I., zu Castell, W., Tsvetkov, P. & Antonov, A. V. AERS spider: an online interactive tool to mine statistical associations in Adverse Event Reporting System. *Pharmacoepidemiol Drug Saf* 23, 795-801 (2014).
103. Böhm, R. *et al.* OpenVigil FDA - Inspection of U.S. American Adverse Drug Events Pharmacovigilance Data and Novel Clinical Applications. *PLoS One* 11, e0157753 (2016).
104. Sarangdhar, M. *et al.* Data mining differential clinical outcomes associated with drug regimens using adverse event reporting data. *Nat Biotechnol* 34, 697-700 (2016).
105. 1 型糖尿病発症に関連する要因の探索. 令和 6 (2024) 年度学部学生による自主研究奨励事業研究成果報告書 <https://hdl.handle.net/11094/101281> (2025).
106. Kunitsu, Y. The Potential of GPT-4 as a Support Tool for Pharmacists: Analytical Study Using the Japanese National Examination for Pharmacists. *JMIR Medical Education* 9, e48452 (2023).
107. Dennstädt, F., Hastings, J., Putora, P. M., Schmerder, M. & Cihoric, N. Implementing large language models in healthcare while balancing control, collaboration, costs and security. *npj Digit. Med.* 8, 1-4 (2025).
108. Wu, Y. *et al.* Embracing ChatGPT for Medical Education: Exploring Its Impact on Doctors and Medical Students. *JMIR Medical Education* 10, e52483 (2024).
109. Mortlock, R. & Lucas, C. Generative artificial intelligence (Gen-AI) in pharmacy education: Utilization and implications for academic integrity: A scoping review. *Exploratory Research in Clinical and Social Pharmacy* 15, 100481 (2024).
110. Sendekie, A. K. *et al.* Artificial intelligence in community pharmacy practice: Pharmacists' perceptions, willingness to utilize, and barriers to implementation. *Explor Res Clin Soc Pharm* 16, 100542 (2024).
111. Naito, H. The GPT-4o Shock Emotional Attachment to AI Models and Its Impact on Regulatory Acceptance: A Cross-Cultural Analysis of the Immediate Transition from GPT-4o to GPT-5. Preprint at <https://doi.org/10.48550/arXiv.2508.16624> (2025).
112. Gangavarapu, A. LLMs: A Promising New Tool for Improving Healthcare in Low-Resource Nations. *2023 IEEE Global Humanitarian Technology Conference (GHTC)* 252-255 (2023)
doi:10.1109/GHTC56179.2023.10354650.
113. Temsah, A. *et al.* DeepSeek in Healthcare: Revealing Opportunities and Steering Challenges of a New Open-Source Artificial Intelligence Frontier. *Cureus* 17, e79221.
114. Jonnagaddala, J. & Wong, Z. S.-Y. Privacy preserving strategies for electronic health records in the era of large language models. *npj Digit. Med.* 8, 1-3 (2025).

115. Zong, H. *et al.* Large Language Models in Worldwide Medical Exams: Platform Development and Comprehensive Analysis. *J Med Internet Res* 26, e66114 (2024).
116. Hu, E. J. *et al.* LoRA: Low-Rank Adaptation of Large Language Models. Preprint at <https://doi.org/10.48550/arXiv.2106.09685> (2021).
117. Japanese National Examination for Pharmacists.
https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryoku/iyakuhin/yakuzaishi-kokkashiken/index.html.
118. SEED, a free study app for the Japanese National Examination for Pharmacists.
<https://yakuzaseed.com>.
119. [2505.16661] A Japanese Language Model and Three New Evaluation Benchmarks for Pharmaceutical NLP. <https://arxiv.org/abs/2505.16661>.
120. mocobeta. janome/CONTRIBUTING.md at master · mocobeta/janome. *GitHub*
<https://github.com/mocobeta/janome/blob/master/CONTRIBUTING.md>.
121. Abdin, M. *et al.* Phi-4 Technical Report. Preprint at <https://doi.org/10.48550/arXiv.2412.08905> (2024).
122. Guo, D. *et al.* DeepSeek-R1 incentivizes reasoning in LLMs through reinforcement learning. *Nature* 645, 633–638 (2025).
123. Qwen *et al.* Qwen2.5 Technical Report. Preprint at <https://doi.org/10.48550/arXiv.2412.15115> (2025).
124. cyberagent/DeepSeek-R1-Distill-Qwen-14B-Japanese · Hugging Face.
<https://huggingface.co/cyberagent/DeepSeek-R1-Distill-Qwen-14B-Japanese>.
125. cyberagent/DeepSeek-R1-Distill-Qwen-32B-Japanese · Hugging Face.
<https://huggingface.co/cyberagent/DeepSeek-R1-Distill-Qwen-32B-Japanese>.
126. D3 Center, Osaka University » Blog Archive » SQUID. <https://www.hpc.cmc.osaka-u.ac.jp/en/squid/>.
127. Saito, K. *et al.* Why We Build Local Large Language Models: An Observational Analysis from 35 Japanese and Multilingual LLMs. Preprint at <https://doi.org/10.48550/arXiv.2412.14471> (2024).
128. Kaplan, J. *et al.* Scaling Laws for Neural Language Models. Preprint at <https://doi.org/10.48550/arXiv.2001.08361> (2020).
129. A Survey of Large Language Models in Medicine: Progress, Application, and Challenge.
<https://arxiv.org/html/2311.05112v5>.
130. Dantas, P. V., Sabino da Silva, W., Cordeiro, L. C. & Carvalho, C. B. A comprehensive review of model compression techniques in machine learning. *Appl Intell* 54, 11804–11844 (2024).
131. Fomenko, V., Yu, H., Lee, J., Hsieh, S. & Chen, W. A Note on LoRA. Preprint at <https://doi.org/10.48550/arXiv.2404.05086> (2024).
132. Giuffrè, M., Kresevic, S., Pugliese, N., You, K. & Shung, D. L. Optimizing large language models in digestive disease: strategies and challenges to improve clinical outcomes. *Liver International* 44, 2114–2124 (2024).
133. Alkhalaf, M., Shen, J., Chang, H.-C. (Rita), Deng, C. & Yu, P. Fine-tuning large language models for effective nutrition support in residential aged care: a domain expertise approach. in (2024).
doi:10.1101/2024.07.21.24310775.

134. Nugraha, G. P., Suadaa, L. H. & Pramana, S. Fine-Tuning Large Language Models for Text-to-SQL Tasks in Agricultural Census Anomaly Detection. *2024 International Conference on Electrical Engineering and Informatics (ICELTICs)* 136-140 (2024)
doi:10.1109/ICELTICs62730.2024.10776457.
135. Yu, Y. *et al.* Fine-Tuning Large Language Models to Improve Accuracy and Comprehensibility of Automated Code Review. *ACM Trans. Softw. Eng. Methodol.* 34, 1-26 (2025).
136. Jiang, S., Dai, Y., Jia, H., Wang, Y. & Wang, H. IntelliChain Stars at the Regulations Challenge Task: A Large Language Model for Financial Regulation. in *Proceedings of the Joint Workshop of the 9th Financial Technology and Natural Language Processing (FinNLP), the 6th Financial Narrative Processing (FNP), and the 1st Workshop on Large Language Models for Finance and Legal (LLMFinLegal)* (eds. Chen, C.-C. et al.) 371-384 (Association for Computational Linguistics, Abu Dhabi, UAE, 2025).

主論文

Asano, H. et al. Safety comparisons among monoamine oxidase inhibitors against Parkinson's disease using FDA adverse event reporting system. Sci Rep 13, 19272 (2023).