

別添 1

令和元年度厚生労働行政推進調査事業費補助金

厚生労働科学特別研究事業

医薬品評価情報作成システムの有用性評価研究

令和元年度 総括・分担研究報告書

研究代表者 鹿野 真弓

令和2(2020)年 3月

目 次

I . 総括研究報告

医薬品評価情報作成システムの有用性評価研究 鹿野真弓	-----	1
-------------------------------	-------	---

II . 分担研究報告

1 . 品質分野等検討 鹿野真弓	-----	7
---------------------	-------	---

2 . ADME分野等検討 永井尚美	-----	37
-----------------------	-------	----

3 . 山田 博章 に関する検討	-----	47
------------------	-------	----

III . 研究成果の刊行に関する一覧表	-----	60
----------------------	-------	----

別添 3 (総括) 研究報告書の作成上の留意事項

1. 「A. 研究目的」について
 - ・厚生労働行政の課題との関連性を含めて記入すること。
2. 「B. 研究方法」について
 - (1) 実施経過が分かるように具体的に記入すること。
 - (2) 「(倫理面への配慮)」には、研究対象者に対する人権擁護上の配慮、研究方法による研究対象者に対する不利益、危険性の排除や説明と同意(インフォームド・コンセント)に関わる状況、実験に動物に対する動物愛護上の配慮など、当該研究を行った際に実施した倫理面への配慮の内容及び方法について、具体的に記入すること。倫理面の問題がないと判断した場合には、その旨を記入するとともに必ず理由を明記すること。

なお、ヒトゲノム・遺伝子解析研究に関する倫理指針(平成25年文部科学省・厚生労働省・経済産業省告示第1号)、人を対象とする医学系研究に関する倫理指針(平成26年文部科学省・厚生労働省告示第3号)、遺伝子治療等臨床研究に関する指針(平成31年厚生労働省告示第48号)、厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針(平成18年6月1日付厚生労働省大臣官房厚生科学課長通知)及び申請者が所属する研究機関で定めた倫理規定等を遵守するとともに、あらかじめ当該研究機関の長等の承認、届出、確認等が必要な研究については、研究開始前に所定の手続を行うこと。
3. 「C. 研究結果」について
 - ・当該年度の研究成果が明らかになるように具体的に記入すること。
4. 「F. 健康危険情報」について
 - ・研究分担者や研究協力者の把握した情報・意見等についても研究代表者がとりまとめて総括研究報告書に記入すること。
5. その他
 - (1) 日本工業規格A列4番の用紙を用いること。
 - (2) 文字の大きさは、10～12ポイント程度とする。

令和元年度厚生労働行政推進調査事業費補助金（厚生労働科学特別研究事業）
（総括）研究報告書

医薬品評価情報作成システムの有用性評価研究

研究代表者 鹿野 真弓 東京理科大学薬学部薬学科・教授

研究要旨

新医薬品開発のグローバル化による規制当局間の国際連携が進む中、日本の医薬品評価結果の情報を迅速に世界に発信する必要がある。本研究においては、PMDAの新医薬品の審査報告書の英語版作成におけるAI自動翻訳ツールによる翻訳の有用性を確認し、その活用にむけた検討を行った。審査報告書のAI翻訳ツールによる翻訳産物について、専門性の観点及び英語の正確性の観点から評価した。その結果、翻訳不備の要因がいくつか特定され、原文作成の段階で翻訳不備を防ぐ方策を取るとともに、定型記載や専門用語の学習・カスタマイズによるAI自動翻訳ツールの性能向上や事後的な人手によるチェックが必要と考えられた。しかし、AI自動翻訳特有の翻訳不備の一部は特段のパターンによらず突発的に発生することから、そのチェックを行う際には審査報告書全体について原文と英文の比較による確認作業が必要となり、人的リソースの軽減効果は限定的で、審査報告書英訳の公表版作成に活用することは現時点では現実的ではないと考えられる。一方で、各分野の専門性を有する者であれば、翻訳不備があっても論旨を把握することは可能と考えられ、例えば規制当局間で早急な情報提供を求められる場合など、必要に応じてAI自動翻訳ツールで英訳を行い、一定の翻訳不備が含まれることを説明した上で最低限の確認・修正のみを行った審査報告書のAI自動翻訳英文版を情報提供することが有用なケースはあると考えられる。AI自動翻訳ツールは、学習やカスタマイズ等によるレベルアップのみならず、翻訳エンジン自体の性能向上も期待されるものであることから、今後、より質の高い自動翻訳ツールが使用可能となることが望まれる。

研究分担者氏名

永井 尚美 武蔵野大学薬学部・教授
山田 博章 横浜薬科大学臨床薬学科・教授

A．研究目的

医薬品・医療機器等の開発・流通のグローバル化や製品の多様化を受けて、関連規制制度は複雑かつ高度化しており、国際的な協力の重要性が世界の規制当局間の共通認識となりつつある。こうした中、世界保健機関（WHO）や、各国/地域の医薬品規制当局のリーダーで構成される薬事規制当局国際連携組織（ICMRA）では、情報共有を視野に入れて各規制当局のリソースをより有効に活用する取組みが進められている。

さらに本邦においては、日本で承認された優れた医薬品・医療機器がアジア諸国で受け入れられ、迅速に患者に提供されることを目指しており、我が国の医薬品等の評価結果や安全性情報の

英語等での発信を積極的に行っていく必要がある。規制当局が専門的内容を含む評価結果の英語版を正確かつ迅速に作成する方策として、AIを活用した自動翻訳ツールの導入が考えられる。

AI自動翻訳ツールの性能は、近年、著しく進歩しており、本研究では、代表的なAI自動翻訳ツールを用いて自動翻訳の正確性等の特徴や課題を評価することでその有用性を確認し、規制当局の医薬品評価情報作成のパフォーマンス向上に資する活用法に関する研究を行うことを目的とする

B．研究方法

1) 自動翻訳ツールの選定

医学・薬学の専門用語にも十分に対応できるAI等を活用した自動翻訳ツールを利用するため、本学HP上で研究への協力を広く呼びかけた。応募された4社のうち、本研究の主旨や本研究事業の仕組み

を踏まえて協力を了承された以下のツールを用いることとした。

- T-400 (Translation For Onsha Only : 株式会社ロゼッタ)
- COTOHA (株式会社みらい翻訳/NTTコミュニケーションズ/株式会社翻訳センター)。

これらの自動翻訳ツールについては、いずれも独立行政法人医薬品医療機器総合機構(以下、PMDA)が、その英訳とともに公表している新医薬品の承認審査報告書や医薬品規制調和国際会議(以下、ICH)各種ガイドライン等の薬事規制関連文書を用いて学習を行ったものを利用した。

なお、本研究においては2つの自動翻訳ツールの性能比較が目的ではないため、自動翻訳ツール名を【A】あるいは【B】と記載する。

2) 評価対象文書の選定

独立行政法人医薬品医療機器総合機構(以下、PMDA)が公表している新医薬品の審査報告書を対象に用いることとした。平成30年以降に承認された新有効成分含有医薬品のうち、効能・効果に偏りが無いよう以下の4つの審査報告書を選定した。

- イミフィンジ点滴静注 120mg、同点滴静注 500mg (以下、イミフィンジ) 抗悪性腫瘍薬
- スキリージ皮下注 75mg シリンジ 0.83mL (以下、スキリージ): 炎症性疾患治療薬
- ヘムライブラ皮下注 30mg、同 60mg、同 90mg、同 105mg、同 150mg (以下、ヘムライブラ): 血液疾患治療薬
- レルミナ錠 40mg (以下、レルミナ): 婦人科疾患治療薬

詳細は、別添4の研究代表者 鹿野の報告書を参照されたい。

3) 評価方法

上記で選択した4つの審査報告書について、自動翻訳ツール【A】及び【B】を用いて翻訳した翻訳産物のうち、品質分野、毒性分野、臨床分野の評価、及び英語の専門性の観点からの評価は、研究代表者の鹿野が研究協力者の参加も得て実施した。ADME分野については研究分担者の永井が、薬理分野につ

いては研究分担者の山田が評価を行った。

評価スコア含め評価方法の詳細については、別添4の研究代表者、各研究分担者の報告書を参照されたい。

(倫理面への配慮)

自動翻訳ツールの採用の経緯ならびに評価方法及び結果を本報告書で公表し透明性を確保することで利益相反が無いことを示す。

また本研究は、再生医療等の安全性の確保等に関する法律、人を対象とする医学系研究に関する倫理指針、ヒトゲノム・遺伝子解析研究に関する倫理指針、動物実験等の実施に関する基本指針、省令GCP、臨床研究法の対象に該当しない。

C. 研究結果

本項においては、各専門分野に特徴的に見られた翻訳不備を中心に記載する。専門分野に関係なく確認された翻訳不備については「D.考察 2) 全体的に見られた翻訳不備の整理」の項でまとめて記載する。

1) 品質分野の英訳文書の評価

4つの品目の審査報告書の翻訳産物のうち品質分野について、修正しなくても意味が通ると評価されたスコア2及び3の文の割合は、翻訳ツール【A】の産物の平均が89.2%、【B】の産物の平均が86.8%と、概ね高い質の翻訳産物が得られていた。

翻訳ツール【A】の産物は、マスキング箇所の前後で構文が不正確になる、あるいは訳漏れが発生するケースが多いことが確認された。ICH品質ガイドラインで規定されている専門用語は概ね正しく翻訳されていた。翻訳ツール【B】の翻訳産物では、原文の一部がマスキングされていると一文全体をマスキングして翻訳する箇所が散見された。また、ICH品質ガイドラインで規定されている専門用語や公定書名等を正しく翻訳されない場合が多く見られた。

2) 薬理分野の英訳文書の評価

審査報告書の非臨床薬理分野を構成する全センテンスの全角換算文字数は、イミフィンジ、スキリージ及びヘムライブラでは、25文字超、50文字以下のセンテンスがピークで約35%を占めていた。レルミナは、25文字以下のセンテンスが約30%で最も多く、150文字超の長文も含んでいた。

翻訳英文の評価スコア分布については、全体的に、修正しなくとも情報として意味が通じるスコア2以上の英文が約80%を占めていた。イミフィンジ、スキリージ及びヘムライブラでは、【B】に比較して【A】の方が、スコア3の割合が高くスコア0の割合が低かった。レルミナでは逆に【B】の方がスコア3の割合が高くスコア0の割合が低かった。

3) ADME分野の英訳文書の評価

審査報告書の解釈及び翻訳版での文意伝達に関して、専門的立場からは修正しなくても意味が通ると判断されたスコア2及び3の箇所割合は、67～94%の範囲であり、全188箇所における平均は、【A】及び【B】の翻訳産物で83%及び85%であった。同一品目においては、翻訳ツール間の顕著な差は認められなかった。品目によってスコア2及び3の割合に差が生じた背景として、ADME評価情報として選定した非臨床薬物動態、臨床薬理及び審査内容が品目ごとに異なるほか、疾患領域や品目特性が影響すると考えられた。

翻訳不備(スコア0及び1)と判断した事由のうち、ADME分野に特徴的なものとしては、一部の専門用語の選択のほか、健康被験者とすべき箇所を患者、表中の項目において、非臨床薬物動態データであるにも関わらずヒトのデータとして取り扱われている箇所があった。

4) 毒性分野の英訳文書の評価

4つの品目の審査報告書の翻訳産物のうち、毒性分野について、修正しなくても意味が通ると評価されたスコア2及び3の割合は、スキリージのみ翻訳ツール【A】【B】の翻訳産物とも50%程度、ヘムライブラはいずれも85%程度、それ以外の2品目は、いずれの翻訳産物も約75%前後であった。

スキリージの【A】【B】のいずれの翻訳産物とも評価が低かった原因は、いずれも原文にない表現や用語が自動的に追加された箇所が複数見られたことにあった。

また翻訳ツールにかかわらず「不適切用語」が散見され、毒性分野で慣用される表現が正しく翻訳されていないケースが多かった。

5) 臨床分野の英訳文書の評価

4つの品目の審査報告書の臨床分野の評価は、品目ごとに異なる評価者が担当したため評価の視点が異なる可能性が考えられ、品目間の差異を比較することは適切ではないが、同じ品目の【A】と【B】の翻訳産物の比較は可能と考えた。修正しなくても意味が通ると評価されたスコア2及び3の文の割合は、イミフィンジ、スキリージ、ヘムライブラは【A】の産物の方が【B】に比較して高い傾向が見られた。一方で、レルミナは【A】が【B】より明らかに低い数値であった。

レルミナについては、原文が長文で構成が複雑あるいは不明瞭なものが多く、【A】と【B】の翻訳産物とも重大な訳漏れや英訳の構文の乱れによる誤訳が多数発生していたが、特に翻訳ツール【A】は、脚注番号が挿入されている場合に不適切な箇所を分割して翻訳するため、重大な誤訳につながった箇所が多いことが確認された。

臨床分野に特徴的と思われる翻訳不備はほとんど見られなかったが、他の診療科の薬物に比較して自動翻訳ツールの学習機会が少ないと推察される婦人科領域の専門用語の誤訳が散見された。

6) 英語の正確さの視点での英訳文書の評価

英語の専門性を有する評価者により、英語の正確性の視点から評価が行われた。

特徴的な翻訳不備としては、審査報告書の冒頭にある「審査報告書」の定型記載の翻訳不備が多く、【A】の翻訳産物では販売名のスペルミス、複数の定型的項目名の誤訳の他、糖鎖構造の表記の訳漏れが見られた。【B】の翻訳産物でも販売名のスペルミス、複数の定型的項目名の誤訳が見られた他、通知引用の記載、審査部名の訳も不適切であった。

記号や数字が多数含まれる糖鎖構造の説明では、訳漏れが複数発生していた。また、目次の記載は翻訳されず日本語のままであった。

D. 考察

1) 各分野等に特徴的な翻訳不備の整理

定型記載について

審査報告書冒頭の「審査報告書」には定型的な記載が多いが、翻訳ツール【A】【B】ともに定型的な項目名の誤訳が複数見られ、これらは目次が日本語のままの不備とともに学習あるいはカスタマイズで対応可能と考えられる。一方で、販売名や一般名のスペルミス、有効成分の構造の記載に含まれる記号や成分名称等の誤訳については個別に対応する必要があると考えられた。

品質分野について

品質分野はマスキングされる件数が多くその範囲も長いので、翻訳ツールにかかわらずマスキング箇所の前後での英訳文の構文の乱れや訳漏れが多く発生する。品質専門用語や表現に関しては、特に翻訳ツール【B】はICH品質ガイドラインで規定されている専門用語を適切に学習させることが必要と考えられた。

品質分野の審査報告書は比較的シンプルな文が多いので、マスキング箇所の前後の重点的なチェック及び専門用語・表現の学習が適切になされれば、重大な誤解を生じるリスクはかなり低減できると考える

薬理分野について

薬理分野に特徴的と言える翻訳不備は特段見られなかった。

ADME分野について

健康被験者とすべき箇所が患者として取り扱われている箇所については、文中に引用されている試験名と対応する訳出であると考えられた。また、ADMEは生体内挙動に関わる学問領域であり、審査報告書中では、唯一非臨床と臨床の両領域を含む。ADME特性の記述に加えて、他の非臨床又は臨床分野に言及する場合もあることが影響する可能性が考えられた。

毒性分野について

スキリージにおいて、項目タイトル(翻訳ツール【A】)や表タイトル(翻訳ツール【B】)に関係ない用語が自動的に追加されていたが、これは翻訳ツールが過去の審査報告書の翻訳の際に学習した内容がスキリージの内容と類似していたため、自動的に学習した内容を反映した可能性が考えられる。このようなAI自動翻訳ツールの特徴的機能に由来すると考えられる誤訳を防ぐ方策については、自動翻訳ツールの開発販売会社との相談・検討が必要であろう。

臨床分野について

翻訳ツールの学習機会が少ないと推察される診療科の専門用語に不備が見られた。対応する診療科のガイドラインを学習させる等の対応が考えられる。

なお今回、自動翻訳ツール【A】【B】いずれも薬事に対応したバージョンを使用した。それぞれ品質分野、毒性分野等の各分野に対応したバージョンも用意されているため、審査報告書の各分野ごとに対応するバージョンの自動翻訳ツールを使用すれば、専門用語や分野に特徴的な表現に適切に対応された翻訳産物が得られる可能性が期待される。

2) 全体的に見られた翻訳不備の整理

原文の構成に由来する翻訳不備

複雑で長い文は誤訳を生じやすく、特に自動翻訳ツール【A】は数字、記号、カッコ、脚注番号等を含むと不適切な個所で分割して訳す例が散見された。

対処法としては、自動翻訳前の作業(Pre-edit)として、長文は避けて複数の文を接続詞でつなぐ記載を心がけ、主要なフレーズ・文に主語を明記する等が、自動翻訳後の作業(Post-edit)として数字、記号等の箇所のチェックが考えられる。

記号、数字、カッコ等に由来する翻訳不備

数字、記号、カッコの多用、多重カッコは訳漏れや数字の入れ替わりが発生しやすい。長文に

含まれる脚注番号が別の文に移動する等が発生する。

対処法としては、自動翻訳前の作業(Pre-edit)としてカッコの整理や丸数字の排除の他、数字、記号、脚注番号について人手でチェックする自動翻訳後の作業(Post-edit)が考えられる。

マスキングに由来する翻訳不備

【A】ではマスキングの前後で構文の誤解釈が発生しやすく、【B】では文の1/3程度がマスキングされると全文をマスキングとして翻訳される場合が多い。

対処法としては、マスキングの前後の訳を重点的に人手でチェックする自動翻訳後の作業(Post-edit)が考えられる。

定型記載に関連する翻訳不備

「本薬」に該当する訳は【A】で「this drug」「the this drug」、【B】では「drug」と訳されわかりにくい。自動翻訳前の作業(Pre-edit)として原文で成分記号等に置き換えておく等の対応が考えられる。

「審査報告書」、審査報告(1)(2)の冒頭部分等の定型記載の誤訳、最後に添付された略語表も記載の工夫が必要。

対処法としては、自動翻訳ツールに学習させる、カスタマイズする等の自動翻訳ツールの機能改善・工夫が考えられる。

訳語の不安定さ

一般的な用語について異なる翻訳がなされる場合が散見され、異なる意味に捉えられて誤解を生じる危険性が否定できない。

訳ブレの発生に一定の規則性が見られず、対処法については自動翻訳ツールの開発販売会社との相談・検討が必要かもしれない。

E．結論

自動翻訳ツールによる翻訳は、英語の正確性の面で課題は見られたものの、各分野専門性を有する者であれば、自動翻訳ツールの翻訳産物の内容を一定程度理解することは可能と考えられた。翻訳産物の質をさらに向上させるためには、(1) 自動翻

訳前の作業(Pre-edit)を行うことで可能な範囲で誤訳を防ぐこと、(2) 適切に学習させた自動翻訳ツールを使用すること、(3) 自動翻訳後の作業(Post-edit)において、翻訳産物について、マスキング箇所や、カッコ、数字、記号、脚注を含む長文、図の画像など翻訳不備が生じやすい箇所を重点的にチェックすること、等が考えられる。しかしながら、無関係の語句の挿入や原文誤解釈等のAI自動翻訳特有の翻訳不備の人手によるチェックには追加の人的リソースの確保や校正に長時間が必要となる可能性が高く、現時点では効率的とは言い難い。一方で、その専門分野の知識を有する者であれば、全体の概要を把握することは十分可能と考えられる。

国際連携が進む中、一定の翻訳不備を完全には排除できなくても審査報告書英文版を迅速に情報提供することが必要な場合も考えられる。このような場合において、AI自動翻訳ツールで英訳を行い、一定の翻訳不備が含まれることを説明した上で審査報告書英文版を提供することが有用な場合もあると考えられる。AI自動翻訳ツールは、将来的な性能向上が期待されるものであることから、今後、審査報告書の公開版作成への適用が可能となる自動翻訳ツールが使用可能となることが望まれる。

F．健康危険情報

特になし。

G．研究発表

1. 論文発表

本研究に関連した論文発表はない。

2. 学会発表

本研究に関連した学会発表はない。

H．知的財産権の出願・登録状況

(予定を含む。)

1. 特許取得

なし

2. 実用新案登録
なし

3.その他
なし

令和元年度厚生労働行政推進調査事業費補助金（厚生労働科学特別研究事業）
（分担）研究報告書

品質・毒性・臨床分野等検討

研究代表者 鹿野 真弓 東京理科大学薬学部薬学科・教授

研究要旨

新医薬品開発のグローバル化による規制当局間の国際連携が進む中、日本の医薬品評価結果の情報を迅速に世界に発信する必要がある。本研究においては、PMDAの新医薬品の審査報告書の英語版作成におけるAI自動翻訳ツールによる翻訳の有用性を確認し、その活用にむけた検討を行った。

審査報告書のAI翻訳ツールによる翻訳産物について、品質、毒性、臨床の各分野の翻訳の質を、専門性の観点及び英語の正確性の観点から評価した。その結果、翻訳不備の要因がいくつか特定され、原文作成の段階で翻訳不備を防ぐ方策を取るとともに、定型記載や専門用語の学習・カスタマイズによるAI自動翻訳ツールの性能向上や事後的な人手によるチェックが必要と考えられた。しかし、AI自動翻訳特有の翻訳不備の一部は特段のパターンによらず突発的に発生することから、そのチェックを行う際には審査報告書全体について原文と英文の比較による確認作業が必要となり、人的リソースの軽減効果は限定的で、審査報告書英訳の公表版作成に活用することは現時点では現実的ではないと考えられる。一方で、各分野の専門性を有する者であれば、翻訳不備があっても論旨を把握することは可能と考えられ、例えば規制当局間で早急な情報提供を求められる場合など、必要に応じてAI自動翻訳ツールで英訳を行い、一定の翻訳不備が含まれることを説明した上で最低限の確認・修正のみを行った審査報告書のAI自動翻訳英文版を情報提供することが有用なケースはあると考えられる。AI自動翻訳ツールは、学習やカスタマイズ等によるレベルアップのみならず、翻訳エンジン自体の性能向上も期待されるものであることから、今後、より質の高い自動翻訳ツールが使用可能となることが望まれる。

A．研究目的

医薬品・医療機器等の開発・流通のグローバル化や製品の多様化を受けて、関連規制制度は複雑かつ高度化しており、国際的な協力の重要性が世界の規制当局間の共通認識となりつつある。こうした中、世界保健機関（WHO）や、各国/地域の医薬品規制当局のリーダーで構成される薬事規制当局国際連携組織（ICMRA）では、情報共有を視野に入れて各規制当局のリソースをより有効に活用する取組みが進められている。

さらに本邦においては、日本で承認された優れた医薬品・医療機器がアジア諸国で受け入れられ、迅速に患者に提供されることを目指しており、我が国の医薬品等の評価結果や安全性情報の英語等での発信を積極的に行っていく必要がある。規制当局が専門的内容を含む評価結果の英語

版を正確かつ迅速に作成する方策として、AIを活用した自動翻訳ツールの導入が考えられる。

AI自動翻訳ツールの性能は、近年、著しく進歩しており、本研究では、代表的なAI自動翻訳ツールを用いて自動翻訳の正確性等の特徴や課題を評価することでその有用性を確認し、規制当局の医薬品評価情報作成のパフォーマンス向上に資する活用法に関する研究を行うことを目的とする。

B．研究方法

1) 自動翻訳ツールの選定

医学・薬学の専門用語にも十分に対応できるAI等を活用した自動翻訳ツールを利用するため、本学HP上で研究への協力を広く呼びかけた。応募された4社のうち、本研究の主旨や本研究事業の仕組みを踏まえて協力を了承された以下のツールを用い

ることとした。

- T-400 (Translation For Onsha Only : 株式会社ロゼッタ)
- COTOHA (株式会社みらい翻訳/NTTコミュニケーションズ/株式会社翻訳センター)。

これらの自動翻訳ツールについては、いずれも独立行政法人医薬品医療機器総合機構(以下、PMDA)が、その英訳とともに公表している新医薬品の承認審査報告書や医薬品規制調和国際会議(以下、ICH)各種ガイドライン等の薬事規制関連文書を用いて学習を行ったものを利用した。

なお、本研究においては2つの自動翻訳ツール商品の性能比較が目的ではないため、自動翻訳ツール名を【A】あるいは【B】と記載する。

2) 評価対象文書の選定

独立行政法人医薬品医療機器総合機構(以下、PMDA)が、その英訳とともに公表している新医薬品の審査報告書を対象に用いることとした。平成30年以降に承認された新有効成分含有医薬品のうち、効能・効果に偏りが無いよう以下の4つの審査報告書を選定した。

- イミフィンジ点滴静注 120mg、同点滴静注 500mg (以下、イミフィンジ)
(https://www.pmda.go.jp/drugs/2018/P20180727001/670227000_23000AMX00485_A100_1.pdf) 抗悪性腫瘍剤
- スキリージ皮下注 75mg シリンジ 0.83mL (以下、スキリージ)
(https://www.pmda.go.jp/drugs/2019/P20190329001/112130000_23100AMX00299_A100_1.pdf) 炎症性疾患治療薬
- ヘムライブラ皮下注 30mg、同皮下注 60mg、同皮下注 90mg、同皮下注 105mg、同皮下 150mg (以下、ヘムライブラ)
(https://www.pmda.go.jp/drugs/2018/P20180329003/450045000_23000AMX00450_A100_1.pdf) 血液疾患治療薬
- レルミナ錠 40mg (以下、レルミナ)
(https://www.pmda.go.jp/drugs/2019/P20190124001/400256000_23100AMX00010

_A100_1.pdf) 婦人科疾患治療薬

当初、1次評価としてPMDAのHPで公表されている審査報告書pdf版についてAI自動翻訳ツールを用いて英訳を作成し、詳細な評価に必要な評価項目の探索を試みたが、pdf文書の翻訳産物は著しく精度が低く、適切な評価のためにはwordファイル等から翻訳する必要があることが確認された。そこでPMDAに協力を依頼し、各審査報告書の公表pdf版と同じ内容(pdf版の黒塗り部分はアスタリスク「*」で置き換え)の非公表word文書を自動翻訳ツール【A】あるいは【B】で英訳した文書、並びにそれぞれの英訳文書を同じ翻訳ツールで日本語に逆翻訳した文書を提供頂いた。PMDAから提供された英訳文書を対象に、各分担研究者が評価を行うこととされた。なお、逆翻訳文書は必要に応じて参考とする位置付けであり、本研究における評価対象ではない。

3) 評価方法

日本製薬工業協会の協力を得て、AI自動翻訳ツールを利用している製薬企業各社の担当者から、一般に自動翻訳ツールで良く見られる誤訳のパターン、各社におけるAI自動翻訳ツールの性能評価方法、また翻訳物の利用状況等に関してヒアリングを行った。

その結果も踏まえて、以下の手順で評価を行った。

品質、臨床のそれぞれの分野あるいは全体について、内容的に重要な記載または誤訳が生じやすいと考えられる文を中心に、1~数か所から合計でおおむね30-50文程度を評価対象文として抽出した。誤訳が生じやすいと考えられる文とは、構文が複雑な長文、マスキング箇所、図表、数値、記号等を含む文である。毒性分野については「5. 毒性試験に関する資料及び機構における審査の概略」の項全体を評価対象とした。自動翻訳ツール【A】あるいは【B】を用いて翻訳された文書の評価対象箇所について、1文ごとの評価を行った。

評価は、研究協力者の参加も得て以下のように行った。品質分野、毒性分野については、それぞれの専門性を有する評価者1名が、4つの品目の審査報告書の【A】及び【B】による翻訳産物を評価した。臨床分野については、専門知識を有する評価者4名が、それぞれ1品目ずつの審査報告書の【A】及び【B】による翻訳産物の評価を行った。また、英語の専門性を有する評価者1名が、英語の正確さの観点から4つの品目の審査報告書の任意の箇所を対象に【A】及び【B】による翻訳産物の評価を行った。

評価対象の英訳1文ごとに以下の4段階でスコア化を行った。

スコア3：このままで問題ない

スコア2：完全な英文にするには一部修正が必要だが修正しなくとも意味は通じる

スコア1：一部修正しないと意味が通じない

スコア0：大きな問題があり大幅な修正が必要

スコア0~2で見られた翻訳不備の内容について、「訳漏れ」、「原文の誤解釈」、「訳文の文法/構文」、「不適切用語」、「記号・数字」、「その他」の6種類のいずれに該当するかを記録した。他に、評価者が気づいた特徴等に関して自由記載することとした。

(倫理面への配慮)

自動翻訳ツールの採用の経緯ならびに評価方法及び結果を本報告書で公表し透明性を確保することで利益相反が無いことを示す。

また本研究は、再生医療等の安全性の確保等に関する法律、人を対象とする医学系研究に関する倫理指針、ヒトゲノム・遺伝子解析研究に関する倫理指針、動物実験等の実施に関する基本指針、省令GCP、臨床研究法の対象に該当しない。

C. 研究結果

1) 品質分野の英訳文書の評価

4つの品目の審査報告書の品質分野について自

動翻訳ツール【A】及び【B】により翻訳した産物を評価し、各スコアの個数、評価概要、スコア0あるいは1と評価したおもな理由を表1~表4に示す。また、各スコアの個数をグラフで表したのが、それぞれ図1~図4である。

各翻訳産物について、修正しなくても意味が通ると評価されたスコア2及び3の文の数及びその割合を表5にまとめた。最低で80%、最高で94%の文が2または3と評価され、翻訳ツール【A】の産物の平均が89.2%、【B】の産物の平均が86.8%と、概ね高い質の翻訳産物が得られていた。

翻訳不備の各内容のセンテンス数を表6に示した。翻訳ツール【A】の産物で多く見られた不備は、「不適切用語」「訳文の文法/構文」「訳漏れ」「原文の誤解釈」であった。翻訳ツール【B】の産物で多く見られた不備は、「不適切用語」が多く、他に「訳漏れ」「原文誤解釈」が散見された。

翻訳ツール【A】あるいは【B】の不備について、表1~表4の審査報告書ごとの評価を確認すると、翻訳ツール【A】の産物は、マスキング箇所の前後で構文が不正確になる、あるいは訳漏れが発生するケースが多いことが確認された。また、品質分野で汎用される表現の翻訳が不正確である事例も散見された(例:変化体をchange body、恒常性をhomeostasisと訳。正しくはvariant、consistency。)が、ICH品質ガイドラインで規定されている専門用語は概ね正しく翻訳されていた。翻訳ツール【B】の翻訳産物では、品質分野で汎用される表現のみならず、ICH品質ガイドラインで規定されている専門用語や公定書名等を直訳し、正しい用語が使われていない場合が多く見られた。マスキング部分が多い文は文全体をマスキングした形で翻訳する「訳漏れ」の事例は、翻訳ツール【A】の場合と同様に確認された。

なお、数値で示すことは難しいが、翻訳ツール【B】の翻訳産物は英語として読みやすい印象であった。

2) 毒性分野の英訳文書の評価

4つの品目の審査報告書の毒性分野について自動翻訳ツール【A】及び【B】により翻訳した産物を評価し、各スコアの個数、図表の情報、スコア0

あるいは1と評価したおもな理由を表7～表10に示す。また、各スコアの個数をグラフで表したのが図5～図8である。

各翻訳産物について、修正しなくても意味が通ると評価されたスコア2及び3のセンテンス数及びその割合を表11にまとめた。スキリージのみ【A】【B】の翻訳産物とも50%程度、ヘムライブラはいずれも85%程度、それ以外の2品目は、いずれの翻訳産物も約75%前後であった。平均で翻訳ツール【A】【B】の産物とも75%前後であった。

以上のように品目間のばらつきが大きいことから、個別品目の評価状況を表7～表10に基づいて確認したところ、スキリージでは翻訳ツール【A】【B】いずれの翻訳産物とも原文にない表現や用語が自動的に追加された箇所が複数見られたことが低評価の原因であった。ヘムライブラでは原文が難解なための誤解釈や不適切用語は見られたものの件数は少なかった。

翻訳不備の内容ごとのセンテンス数を表12に示した。ヘムライブラを除く3品目で「原文の誤解釈」の件数が多く、原文が難解な長文であることが原因と考えられた。なおイミフィンジの【A】の翻訳産物で訳漏れが多いが、これも原文に難解な長文が多かったためと推察された。他には、翻訳ツールにかかわらず「不適切用語」が散見され、毒性分野で慣用される表現が正しく翻訳されていないケースが多かった。

3) 臨床分野の英訳文書の評価

4つの品目の審査報告書の臨床分野について自動翻訳ツール【A】及び【B】により翻訳した産物を評価し、各スコアの個数、評価概要、スコア0あるいは1と評価したおもな理由を表13～表16に示す。また、各スコアの個数をグラフに示したのが図9～図12である。

各翻訳産物について、修正しなくても意味が通ると評価されたスコア2及び3のセンテンス数及びその割合を表17にまとめた。臨床分野の評価は、品目ごとに異なる評価者が担当したため評価の視点が異なる可能性が考えられ、品目間の差異を比較することは適切ではないが、同じ品目の【A】

と【B】の翻訳産物の比較は可能と考えた。イミフィンジ、スキリージ、ヘムライブラは【A】の産物の方が【B】に比較して2あるいは3と評価されたセンテンスの割合が高い傾向が見られた。一方で、レルミナは【A】が【B】より明らかに低い数値であった。

そこで、レルミナの評価状況を表16に基づいて確認したところ、原文が長文で構成が複雑あるいは不明瞭なものが多く、重大な訳漏れや英訳の構文の乱れによる誤訳が多数発生していた。特に翻訳ツール【A】は、引用先の脚注番号が挿入されている場合に不適切な個所で文を分割して翻訳するため、重大な誤訳につながった箇所が多いことが確認された。【A】では他に、「the this drug group」との訳が2か所あった。この誤訳は他の品目、他の分野の翻訳でも頻出している。【B】の産物で上記以外の誤訳としては、本薬を定冠詞なしの「drug」と訳す、表内の「135例^a」を「^a of 135 subjects」と訳すなどが見られた。

他の品目の評価において、0または1とされたおもな不備は以下の通りである。

イミフィンジでは、【A】の翻訳産物で原文に無い数字の挿入が2か所で確認された他、「the this drug group」に加えて「the investigational product this drug group」、項目立ての、を(1)、(b)と訳すことによる項立ての混乱、「それぞれ()55(- 3¹⁴) ~ 406)」が正確に翻訳されず情報が把握できない等、カッコ、記号、数値が混在する表記の訳が不正確な場合などが見られた。【B】の翻訳産物では、定冠詞(the)や一部フレーズの細かい訳漏れが意味に影響する箇所が多数見られた。他に、マスキングや長文記載が原因と考えられる訳の乱れ、図のタイトルの語順の乱れやpdf画像を貼り付けたと思われる図内の日本語が一部しか訳されていない等が確認された。

スキリージでは、翻訳ツール【A】の産物で、関係ない薬剤名が挿入されていた。文中に脚注番号や図表番号があるとその後で訳の構文の乱れが多数発生していた他、同一用語の訳のブレ、カッコ内の数値が不正確な箇所があった。不適切用語の訳も確認され、「目的」を「EYES Common」とする

突拍子もない訳も見られた。他に、原因となるパターンが不明の訳漏れも散見された。【B】の産物では脚注文頭に不要な「and」の挿入や原文が長文であることが原因と推定される不自然な訳が複数確認された。他に、表中の数字「1,971例」を「11971 subjects」とする誤訳、薬事上、重要な固有名詞「リスク管理計画」の誤訳も見られた。

ヘムライブラでは【A】の産物で無関係の薬剤名の挿入が確認された。他に、同一用語の翻訳のブレや、多重カッコを用いた症例数の説明の誤訳、並列して列挙する記載や表番号を引用した際の訳の乱れが見られた。一方で、長文を複数に分割して訳されたことで原文よりわかりやすい記載となっている箇所も複数あった。【B】の産物では、【A】と同じ箇所でも多重カッコを用いた症例数の説明の誤訳が生じていた。また、4群の観察期間の説明において1群だけ異なる説明があるにもかかわらず他群と同じ訳にされていた。他に、文頭に数字が来る濃度記載があった場合に数字のみ訳してその文節を訳さない、「7例4件」を「7 subjects」と訳、本薬を定冠詞なしの「drug」と訳等の不備が見られた。

表18では、内容への影響とは関係なく確認された不備の情報を見ることができる。レルミナ以外の3品目では【B】で不備の合計が【A】より多い傾向が見られ、特に不適切用語が【A】より多い傾向が見られた。

4) 英語の正確さの視点での英訳文書の評価

英語の専門性を有する評価者により、英語の正確性の視点から評価が行われた。

4つの品目の審査報告書について自動翻訳ツール【A】及び【B】により翻訳した産物を評価し、各スコアの個数、スコア0あるいは1と評価したおもな理由を表19～表22に示す。また、各スコアの個数をグラフに示したのが図13～図16である。

各翻訳産物について、修正しなくても意味が通ると評価されたスコア2及び3のセンテンス数及びその割合を表23に、英訳不備内容の件数を表24にまとめた。表23からは、各専門性を有する評価者による評価(表5、表11、表17)と比較し

て修正が必要と評価されたセンテンスの割合が明かに高いことが確認された。また表24では、各専門性を有する評価者による評価(表6、表12、表18)と比較して「その他」と分類された不備が多かった。

各個別品目の評価について、各分野の専門性を有する評価者での評価で見られなかった事項を中心に以下に記載する。

イミフィンジについては、審査報告書の非臨床薬理及び臨床の項、ならびに審査報告(2)から評価対象文が抽出された。自動翻訳ツール【A】の翻訳産物では、英語で使用されない丸数字(、等)を含む文において、単なる数字として表記され、「0.1又は1mg/kg」を「10.1 or 21 mg/kg」と誤った数値が表示されていた。また、審査報告(2)の冒頭部分の記載については、販売名のイミフィンジ「Imifinzi」を「Imifindi」と誤ったスペルで記載されていた。【B】の翻訳産物では、「本薬の用法・用量について」を「Dosage and administration of ramucirumab」と関係ない薬剤名を挿入して翻訳されていた。また、販売名のみでなく一般名「durvalumab」も「dulvalumab」と誤ったスペルで記載されていた。

スキリージについては、審査報告書の臨床薬理及び臨床の項から評価対象文が抽出された。毒性分野、臨床分野で見られた不備と同様のもの以外の特徴的不備は見られなかった。具体的には、【A】の翻訳産物で複数投与群の説明で用量と症例数の数値や群の名称の番号等、数字が多く出現する文では、数字の入れ替わりが多数生じていた。また、数値とカッコが続く文で訳漏れが発生、脚注番号の前後で文を分割したため文法的に不完全な訳となったケース、グラフ画像の日本語が翻訳されない、等が確認された。【B】の翻訳産物では、マスキングとカッコが重なる部分、カッコや数値を含む長文で構文の乱れや訳漏れが発生していた。また、「薬物動態」、「血漿中薬物濃度データ」、「皮下投与」、「抗体価」の訳が不適切であった。

ヘムライブラについては、審査報告書の冒頭にある「審査報告書」及び品質の項から評価対象文が抽出された。「審査報告書」では定型的な記載が多

いが、定型的な用語が正確に翻訳されていない箇所が多かった。【A】の翻訳産物では販売名HemlibraをHemribulaと誤ったスペルで記載されていた他、[目次]を[Eyes] Next、「本質」を「References」、「構造」を「Design」とする誤訳が見られた。また、糖鎖構造の記述では「Glc」「GlcNAc」の訳漏れが見られた。【B】の翻訳産物でも「ヘムライブラ」を「Heme Library」、「本質」を「Essential」、「構造」を「Preparation」とする誤訳が見られた他、通知引用の記載、審査部名の訳も不適切であった。記号や数字が多数含まれる糖鎖構造の説明では、訳漏れが複数発生していた。また、目次の記載は翻訳されず日本語のままであった。

レルミナについては、毒性及び臨床の項、審査報告(2)ならびに略語表から評価対象文が抽出された。毒性分野、臨床分野で見られた不備と同様、原文が不適切な構成のため誤解釈や訳漏れが発生する等が見られた。略語表については、細かい記載整備が必要な箇所が多数確認された。

D. 考察

1) 品質分野の英訳文書の評価について

評価した医薬品のうち3品目が遺伝子組み換え医薬品、1品目が化学合成医薬品であった。一般に、遺伝子組み換え医薬品の製造方法や品質管理は化学合成医薬品のそれよりも記載が複雑になる場合が多いが、翻訳産物の質に対する影響は特段見られなかった。

品質分野の審査報告書原文の特徴としては、以下があげられる。

- 他の分野に比べると構成が複雑な長文は少なく、比較的シンプルな構文が多い。
- 医薬品の製造や品質管理は各社の知財やノウハウに関する内容を含むため、他の分野に比べてマスキングされる箇所の長さや件数が多い。
- 同じ日本語でも専門分野によって対応する英語が異なる（例：含量は臨床分野ではpotencyを用いる場合があるが品質ではcontent。）。また品質分野で汎用される表

現や言い回しがある。

これらの特徴が、まさに品質分野の自動翻訳ツールでの翻訳に大きく影響を及ぼすと考えられる。

具体的には、2点目に関して、翻訳ツールにかかわらずマスキング箇所を挟むと英訳文の構文が不正確になる場合や、前後の記載の訳漏れを招く場合が多い。翻訳産物について、マスキング箇所の前後を重点的にチェックすることで翻訳文書の質の向上を図ることが可能と考えらる。

3点目の品質専門用語や表現に関しては、翻訳ツール【B】は、ICH品質ガイドラインで規定されている専門用語を適切に学習させることが必要と思われる。品質分野で汎用される表現も、種類はさほど多くないため、いずれの翻訳ツールについても学習・カスタマイズを進めることで、より正確な英訳文書を得ることが期待される。

すなわち、これらの点に対応できれば、1点目に記載したように品質分野の審査報告書は比較的シンプルな文が多いため、海外規制当局等の関係者が読んだ場合に重大な誤解を生じるリスクはかなり低減できるものと考えられる。

2) 毒性分野の英訳文書の評価について

毒性分野の審査報告書ではマスキングは極めてまれであるため、その影響は考慮する必要は無い。一方で、実施される試験の種類や試験法の大枠はICHガイドラインで定められており、品目間で共通あるいは類似した試験が実施される場合が多い。今回、スキリージにおいて、項目タイトル(翻訳ツール【A】)や表タイトル(翻訳ツール【B】)に関係ない用語が自動的に追加されていたが、これは翻訳ツールが過去の審査報告書の翻訳の際に学習した内容がスキリージの内容と類似していたため、自動的に学習した内容を反映した可能性が考えられる。このような誤訳を防ぐ方策については、自動翻訳ツールの開発販売会社との相談・検討が必要であろう。

他に翻訳に影響を及ぼす重要な因子として、原文が長文で日本語として適切な構文となっていない場合が多々見られされた。このような場合、翻訳時に原文の誤解釈や訳漏れを招き、翻訳産物の質

の低下につながる。審査報告書作成時には、例えば1文を極力2行以内とする、箇条書きを活用する、定型的な評価を行う場合には記載様式を可能な範囲で統一する等、論理的に整理された記載を意識して審査報告書を作成することが望まれる。また、データを掲載する表形式を可能な範囲で統一することも、安定した翻訳に資する可能性が考えられる。

さらに、専門用語の統一や毒性分野で汎用される表現の学習が必要と考えられる。ICHガイドラインに記載される事項よりもさらに詳細な専門用語（例：発生学、病理学等に関連した専門用語）の統一の必要性については毒性専門家の間でも認識されており、今後、各学会等で辞書作成の活動が進むことが期待される。

なお、数値化は難しいが、【A】は直訳に近い翻訳が多いため、原文が曖昧な日本語だと誤訳が生じやすい。【B】は学習を踏まえた意識が多いものの、適切でない意識となる場合も見られた。

3) 臨床分野の英訳文書の評価について

臨床分野は、品質分野ほどではないがマスキング箇所が複数あり、それが原因と考えられる誤訳が少数ながら発生していた。

品質分野、毒性分野で見られたような、分野に特徴的な用語や言い回しに関する誤訳は【A】【B】ともに少なかった。これは、臨床分野では病名、症状名等の辞書が充実していることに加え、自動翻訳ツールが臨床関連文書の翻訳に使用されるケースが他の分野の文書に比べて多く学習量も多いためではないかと推察された。

毒性分野と同様に、無関係な数字や薬品名等が挿入された例が【A】【B】でともに確認された。やはり翻訳ツールが過去の審査報告書の翻訳の際に学習した内容が自動的に反映された可能性が考えられる。このような突発的な無関係の数字や単語の挿入は予測が難しい。また、定冠詞等の細かい訳漏れが内容解釈に影響するケースも散見された。「本薬」を「drug」と訳される場合については、原文の段階で成分記号に置き換えることで誤訳を防ぐことが可能と考えられる他、学習での対応も可

能かもしれない。しかし、不規則に生じる細かい訳漏れへの対応は困難と予測され、突発的な挿入と同様に、回避策については翻訳ツールの開発販売会社との相談・検討が必要であろう。

他にも【A】での「the this drug group」など、出現が予測される誤訳は事後に検出して修正することが可能だが、翻訳ツールの学習で対応可能かもしれない。

また、やはり原文が長文で日本語として適切な構文となっていない場合、文の途中にカッコで参照先等を記載する場合や多重カッコを用いる場合に重大な誤訳が生じやすいことも確認された。審査報告書を作成する段階で、可能な範囲で誤訳を生じやすい記載を避けるためのルールを設けることは、論理的に整理された正確な記載を目指すことにもつながり、日本語の審査報告書もより理解しやすいものになると期待される。

なお表17に示したように、全体に【A】の方が【B】より正確に訳されており、用語の適切性や原文の構成が複雑でない場合の訳文構成の面で優れているが、原文の構成が複雑で整理が不十分な点があったレルミナの場合は【B】はより柔軟に対応可能であることが確認された。また、【B】では数字、太字、スペース、フォント等でより優れていた。

4) 英語の正確さの視点での英訳文書の評価について

英語の専門性を有する評価者による評価においては、「審査報告書」、審査報告(2)の冒頭部分等、定型的な記載についての翻訳不備が多数確認された。これらについては自動翻訳ツールの学習やカスタマイズ等によりかなりの部分に対応可能と考えられる。それ以外の箇所で確認された翻訳の不備のほとんどは他の評価者による評価でも指摘されているが、英語の専門性を有する評価者による評価スコアは、分野専門性を有する評価者の評価スコアより全体に低いものであった。分野専門性を有する評価者が評価する際には、英語の正確性に多少欠けていても当該分野の専門知識や前後の記載内容に基づいて欠けた部分を補足しながら内容を理解することが可能かという視点も加わるた

めであろう。極端に言えば、全く意味不明な一文があってもその前後を読むことで全体としてどのようなことが記載されているかを理解できれば大きな問題となる可能性は低い。

さて本研究は、規制当局の医薬品評価情報作成のパフォーマンス向上に資するAI自動翻訳ツール活用法に関する研究を行うものであり、そのためにAI自動翻訳ツールによる翻訳の正確性等の特徴や課題を評価することが重要となる。そのためには、翻訳産物が日本語の原文の意味を忠実に反映し、英語として正確な記載となっているかを指標とした評価を行うことが大前提となる。

その評価を踏まえて課題を整理し、さらに対応法を検討することで翻訳の質の向上を図ることが求められている。

翻訳の質の面での不足を専門知識で補うことはもちろん重要ではあるが、翻訳の不備は内容的な重要性とは関係なく発生しうるため、重要な部分の誤訳が重なると全体の理解にも重大な影響を及ぼすことになる。従って、まずは全体として翻訳の質を向上させ、それでも一定程度発生する翻訳不備について可能な範囲で専門知識でカバーすることを考える必要がある。

5) 確認された翻訳不備の整理

以上の評価において頻出していた翻訳不備の種類及びその対処法を以下に整理した。

原文の構成に由来する翻訳不備

- 複雑で長い文は原文構成の誤解釈を招き、一部の訳漏れ、訳文の不適切な文法/構文を招きやすい。
- 自動翻訳ツール【A】は長文を分割して翻訳するケースが多いが、数字、記号、カッコ、脚注番号等があると不適切な個所で分割されてしまう。
- 原文の構成が不適切なく(日本語としてわかりにくい)文は翻訳不備が起こりやすい。自動翻訳ツールが分割して訳すことで日本語よりわかりやすくなる場合もあるが、稀である。

< 対処法 >

- 自動翻訳前の作業 (Pre-edit)

日本語として理解しやすい文を作成する。長文は避けて複数の文を接続詞でつなぐ記載を心がける。長文や、主語が無い文が複数続く場合などは、主要なフレーズ・文に主語を明記する。

- 自動翻訳後の作業 (Post-edit)

数字、記号、カッコ、脚注番号を排除することは難しいため、自動翻訳ツールでの翻訳後に人手で重点的にチェックする。

記号、数字、カッコ等に由来する翻訳不備

- 数字、記号、カッコ等が続くと訳漏れや数字の入れ替わりが発生しやすい。
- 多重カッコは訳漏れや誤訳を生じやすい。
- 長文に含まれる脚注番号が別の文に移動。
- 脚注番号の片カッコが脱落。
- 【A】ではマスキングの前後で構文の誤解釈が発生しやすい。
- 【B】では文の1/3程度がマスキングされると全文をマスキングとして翻訳されたり、一部日本語が残ったりする。

< 対処法 >

- 自動翻訳前の作業 (Pre-edit)

多重カッコは、() [] 等、形を区別して用いる。丸数字 (、 等) は使用しない。

- 自動翻訳後の作業 (Post-edit)

カッコ内での数字、記号が続く場合は自動翻訳ツールでの翻訳後に人手で重点的にチェックする。

定型記載に関連する翻訳不備

- 「本薬」に該当する訳は【A】で「this drug」「the this drug」、【B】では「drug」と訳されわかりにくい。原文で成分記号等に置き換えておく等の対応が考えられる。
- 「審査報告書」、審査報告(1)(2)の冒頭部分等の定型記載の誤訳が多い。
- 審査報告書の最後に添付された略語表も記載の工夫が必要。

< 対処法 >

- 自動翻訳前の作業 (Pre-edit)

1点目については原文で成分記号等に置き換えておく等の対応が考えられる。

- 自動翻訳ツールの機能改善・工夫

定型記載については自動翻訳ツールに学習させる、

カスタマイズする等が考えられる。

専門用語、専門分野で汎用される表現に関連する記載不備

- 【B】ではICH品質ガイドラインの用語の誤訳が多数発生していた。
- 専門用語に関する訳ブレも見られる。
- 【A】【B】とも、品質分野、毒性分野で慣用的に用いられる表現に対応できない場合があった。
- 学習機会が少ない臨床分野(婦人科領域等)の専門用語の誤訳が発生していた。

<対処法>

- 自動翻訳ツールの機能改善・工夫
- 自動翻訳ツールでは各専門分野に対応したバージョンも使用可能である。今回は薬事に対応したバージョンを用いて翻訳され、臨床分野については他の分野に比べて専門用語の翻訳不備は少なかった。品質、毒性等の他の分野については、各分野に対応したバージョンの翻訳ツールを用いて翻訳することも考えられる。また、専門分野に対応した辞書の作成等も待たれる。なお、臨床用語については自動翻訳ツールの学習機会が多くない診療科の診療ガイドライン等を学習させることも考えられる。

原文にない無関係の語句の挿入

自動翻訳ツールの特徴に由来する翻訳不備であるため、開発販売会社と対応法を相談・検討する必要があると考える。

訳語の不安定さ

専門用語でない一般的な用語についても、異なる翻訳がなされる、いわゆる「訳ブレ」が多く見られた。一文の翻訳として大きな問題は無くても、文によって同じ語が異なる訳とされることで異なる意味に捉えられ、誤解を生じる危険性が否定できない。

<対処法>

訳ブレの発生に一定の規則性が見られないことから、例えば、前後に過去の翻訳事例に類似する表現があった場合にその影響を受ける等、AI自動翻訳の機能に由来する可能性が考えられる。自動翻

訳ツールの開発販売会社との相談・検討が必要な事項かもしれない。

全体を通して

現在もPMDAでは人手で審査報告書の英訳を行い公表しているが、英語版審査報告書の公開までに数か月を要する。迅速性を優先させるために自動翻訳ツールを活用する場合は、前述の対処を行っても一定の翻訳不備は避けることはできず、無関係の語句の挿入や原文誤解釈等のAI自動翻訳特有のエラーの校正や誤訳によるリスクを回避するためには、更なる人的リソースの確保や校正に長期間が必要となる可能性が高い。したがって、自動翻訳産物を審査報告書英訳の公表版に活用することは現実的ではないと考えられる。

一方で、その分野について一定の専門性を有する者であれば、前後の記載や全体の論理展開から概要を理解することは十分可能と思われる。例えば、規制当局間では早急な情報を求められる場合等、必要に応じてAI自動翻訳ツールで英訳を行い、一定の翻訳不備が含まれることを説明した上で、最低限の確認・修正のみを行った審査報告書のAI自動翻訳英文版を情報提供することが有用なケースはあると考えられる。

AI自動翻訳ツールは、学習やカスタマイズ等によるレベルアップのみならず、翻訳エンジン自体の性能向上も期待されており、将来的には審査報告書など公開文書の正式な版への適用が可能となる自動翻訳ツールが使用可能となることが望まれる。

E . 結論

自動翻訳ツールによる翻訳は、英語の正確性の面で課題は見られたものの、各分野専門性を有する者であれば、自動翻訳ツールの翻訳産物の内容を一定程度理解することは可能と考えられた。翻訳産物の質をさらに向上させるためには、(1) 自動翻訳前の作業 (Pre-edit) を行うことで可能な範囲で誤訳を防ぐ、(2) 適切に学習させた自動翻訳ツールを使用すること、(3) 自動翻訳後の作業 (Post-edit) において、翻訳産物について、マスキング箇所や、カッコ、数字、記号、脚注を含む長文、図の画像な

ど翻訳不備が生じやすい箇所を重点的にチェックすること、等が考えられる。しかしながら、無関係の語句の挿入や原文誤解釈等のAI自動翻訳特有の翻訳不備の人手によるチェックには追加の人的リソースの確保や校正に長時間が必要となる可能性が高く、効率的とは言い難い。一方で、その専門分野の知識を有する者であれば、全体の概要を把握することは十分可能と考えられる。

国際連携が進む中、一定の翻訳不備を完全には排除できなくても審査報告書英文版を迅速に情報提供することが必要な場合も考えられる。このような場合において、AI自動翻訳ツールで英訳を行い、一定の翻訳不備が含まれることを説明した上で審査報告書英文版を提供することが有用な場合もあると考えられる。AI自動翻訳ツールは、将来的な性能向上が期待されるものであることから、今後、審査報告書の公開版作成への適用が可能となる自動翻訳ツールが使用可能となることが望まれる。

F．研究発表

1. 論文発表

本研究に関連した論文発表はない。

2. 学会発表

本研究に関連した学会発表はない。

G．知的財産権の出願・登録状況

(予定を含む。)

1. 特許取得

なし

2. 実用新案登録

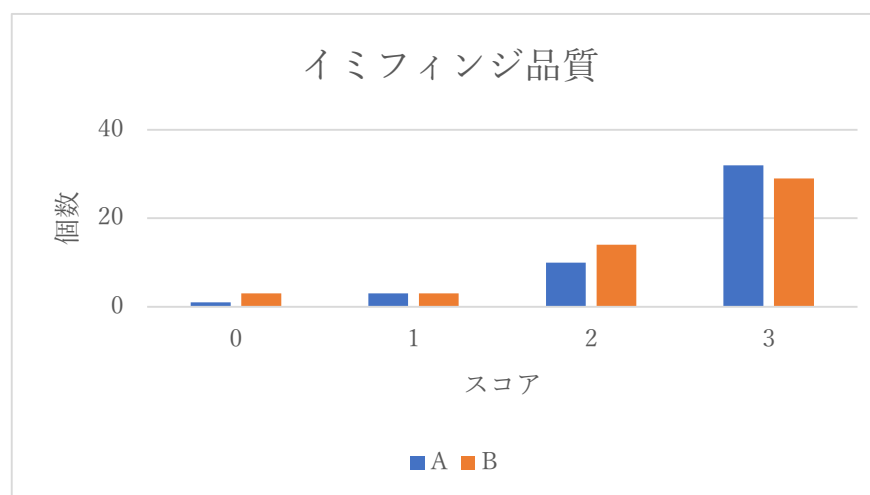
なし

3.その他

【表1】イミフィンジ英訳文書の品質分野評価まとめ

ツ ー ル	スコア				計	概要	0 or 1 のおもな理由
	0	1	2	3			
A	1	3	10	32	46	項目番号と項目名の上にスペースが無く誤解を生じる可能性あり。 マスキング箇所が長い・複数ある場合、構文の解釈を誤って訳出することが多い。	マスキングを挟むことで構文が正しく訳されず意味が変わる。
B	3	3	14	29	49	ICH Q ガイドラインにもある品質専門用語の訳が不安定。学習させれば対応可能か。 マスキング範囲が1文の1/3を超えると全文マスキングで訳出されることが多い。	品質専門用語が正しく訳されないことが多い。 記号を異なる記号に変える場合がある。1文の1/3程度がマスキングされた文が全文マスキングされて訳出。

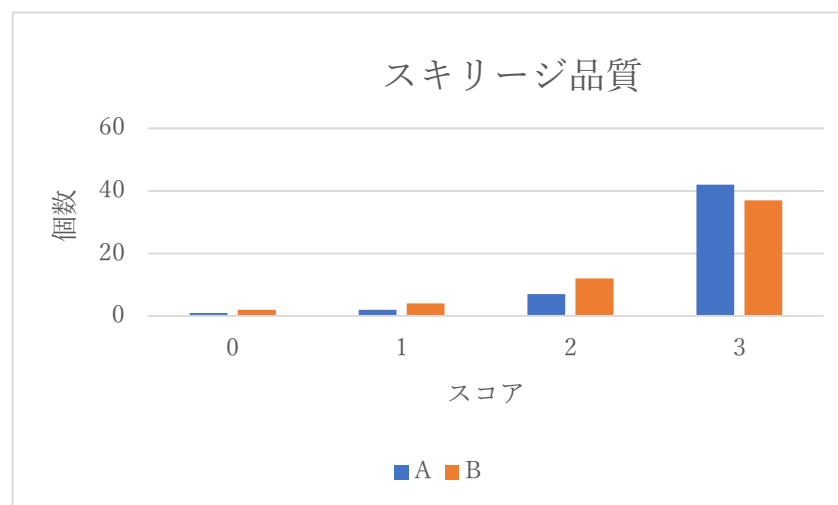
【図1】イミフィンジ英訳文書品質評価スコア分布



【表2】 スキリージ英訳文書の品質分野評価まとめ

ツ ー ル	スコア（個数）				計	概要	0 or 1 のおもな理由
	0	1	2	3			
A	1	2	7	42	52	項目番号と項目名の間スペースが無く、誤解を生じる可能性がある。	非還元を reduction と逆の意味に翻訳。 マスキングされた物質名を別名に置き換えることで不自然なスペースが多数生じ誤解を招く恐れ。
B	2	4	12	37	55	英文自体は A より読みやすい。	品質専門用語の訳が不安定・不正確。構文の誤解釈。 「本薬」を「drug」とのみ訳出。 元文にない参照先を勝手に追加。過去事例の学習のせい。

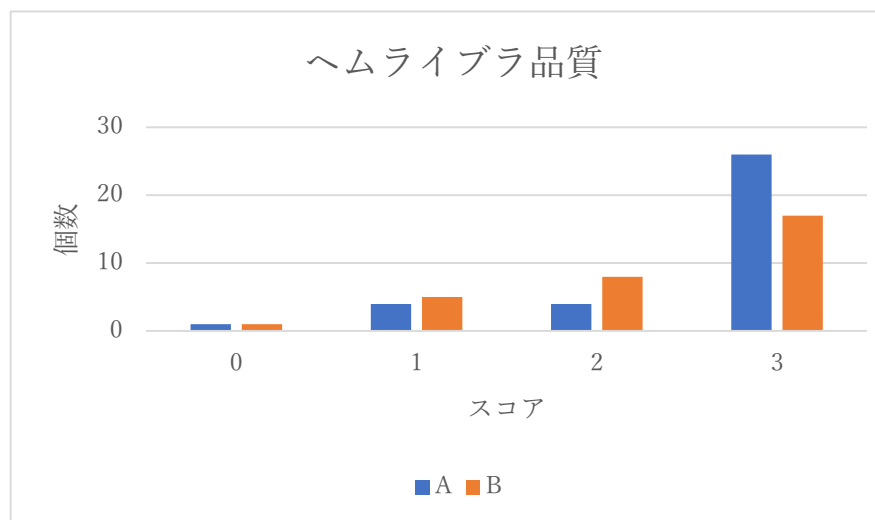
【図2】 スキリージ英訳文書品質評価スコア分布



【表3】 ヘムライブラ英訳文書の品質分野評価まとめ

ツ ー ル	スコア（個数）				計	概要	0 or 1 のおもな理由
	0	1	2	3			
A	1	4	4	26	35	マスク部分が多い箇所は訳漏れ、誤訳。 品質で頻出する用語の訳が不正確。学習で対応可能か。	マスク箇所を含むと構文が不正確、公開部分もマスク。 品質分野で頻出する用語の翻訳が不正確。
B	1	5	8	17	31	意味は通る場合もあるが、ICH ガイドラインで規定されている用語や品質分野で頻出する用語（恒常性、等）を正しく翻訳できない。 カスタマイズすることで対応可能か。	主語、目的語の取り違い。 品質専門用語の翻訳が不安定・不正確。 マスクが多いと日本語のままの箇所あり。

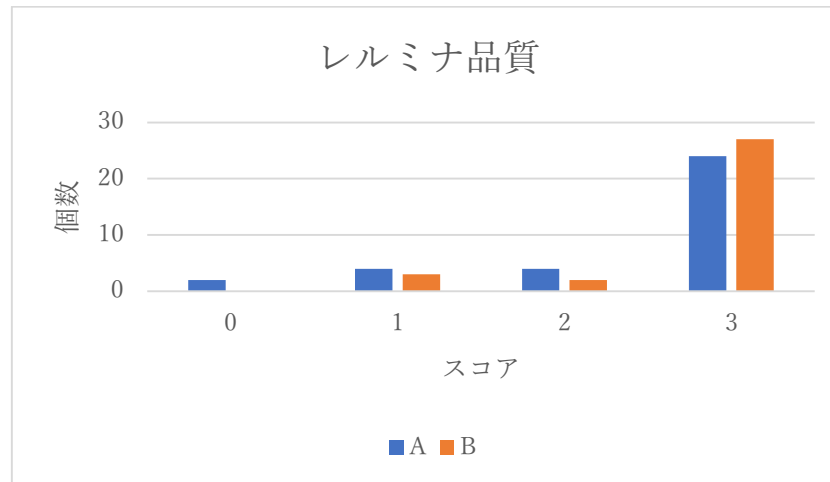
【図3】 ヘムライブラ英訳文書品質分野評価スコア分布



【表4】レルミナ英訳文書の品質分野評価まとめ

ツ ー ル	スコア（個数）				計	概要	0 or 1 のおもな理由
	0	1	2	3			
A	2	4	4	24	34	品質分野で頻出する用語の訳が不正確。	主語の誤り。 品質専門用語の訳が不正確。 マスキング箇所の割合が多いと全文をマスキングして訳出。
B	0	3	2	27	32	品質専門用語、品質分野で頻出する用語の訳が不正確。	品質専門用語の訳が不正確。 マスキング箇所の割合が多いと全文をマスキングして訳出。

【図4】レルミナ英訳文書品質分野評価スコア分布



【表5】各翻訳産物の品質分野で2あるいは3と評価されたセンテンス数及びその割合

審査報告書	イミフィンジ		スキリージ		へムライブラ		レルミナ		合計	
	A	B	A	B	A	B	A	B	A	B
翻訳ツール										
スコア2及び3のセンテンス数	42	42	49	49	30	25	28	29	149	145
評価対象全センテンス数	46	49	52	55	35	31	34	32	167	167
割合(%)	91.3	85.7	94.2	89.1	85.7	80.6	82.4	90.6	89.2	86.8

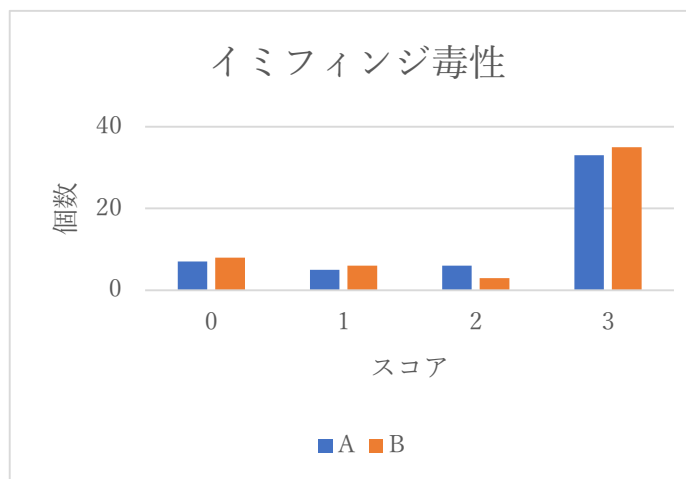
【表6】品質分野の英訳不備内容の件数

対象文書 不備内容	イミフィンジ		スキリージ		へムライブラ		レルミナ	
	A	B	A	B	A	B	A	B
訳漏れ	2	4	2	1	2	2	2	1
原文の誤解釈	0	2	1	2	1	1	3	0
訳文の文法／構文	6	0	1	0	2	0	1	0
不適切用語	6	13	5	12	5	11	4	4
記号・数字	1	2	1	0	0	0	2	0
その他	1	0	1	3	2	0	0	0
合計	16	21	11	18	12	14	12	5

【表7】イミフィンジ英訳文書の毒性分野評価まとめ

ツ ー ル	スコア（個数）				計	図表について	0 or 1 のおもな理由
	0	1	2	3			
A	7	5	6	33	51	【表・図】★M/F は Males/Females に統一？。★Washout は Recovery が正しい。★orbit は正しくない。眼周囲★12 頁の脚注は誤訳。★Table10 の c) が上付きでない。★ELISA は原文にない「処理」注釈が上付きでないのと位置違い。★Table12 試験方法が誤訳（原文が難解）	専門用語、頻出用語が適切に翻訳されていない。 原文が難解で適切に翻訳されない。 「等」が訳されない場合が多い。
B	8	6	3	35	52	【表・脚注】★表 8 「Administration Pathway」は「Dosing Route」が適切（全表）。★表 9 「the rest period」、「Resilient」は「ricovery」が適切。★「of」が不要な箇所あり。★表 10 「Spread-type」は「Enhanced」★注釈の上付き位置、字体大きさ、スペースが不正。★「birth」は「offspring」が適切。★表 11 フォント、上付き位置が不適。★表 12 「in animals」が不要箇所あり。★p16 誤解釈と重複文章あり。	専門用語、頻出用語が適切に翻訳されていない。 原文が難解で適切に翻訳されない。 「胎児／新生児」に対応する数値「30/60」を「30 of 60」と訳出。

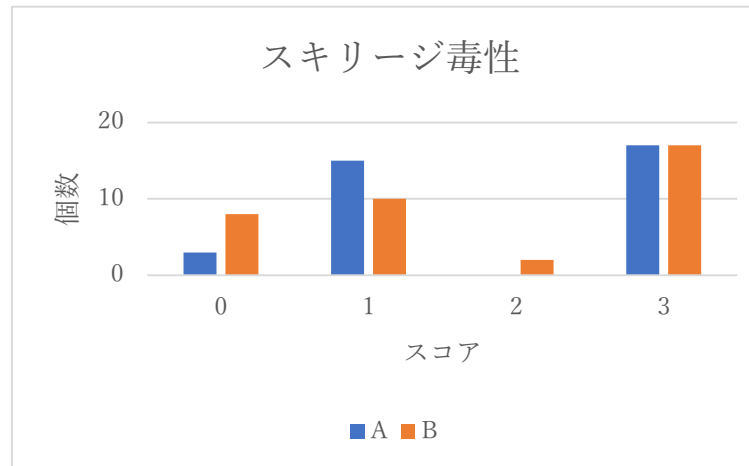
【図5】イミフィンジ英訳文書品質評価スコア分布



【表8】 スキリージ英訳文書の毒性分野評価まとめ

ツ ー ル	スコア (個数)				計	図表について	0 or 1 のおもな理由
	0	1	2	3			
A	3	15	0	17	35	<p>【表全体】★表題の Table と番号間にスペースが必要。★静脈内、皮下投与は略号が通例 (IV、SC 等)。★Washout は Recovery が正しい★上付き記号が無視されている。★表題の大文字・小文字の統一がない。★注釈が違う記号に誤変換。</p>	<p>訳文の構文が不適切。 原文にない内容が自動的に追加された箇所 (項目タイトル) が複数あり。 カッコ記載の訳漏れ。</p>
B	8	10	2	17	37	<p>【表・図】★タイトル末の不要数字。★ポイント数が異なる字体がある。★生殖用語については辞書作成中★審査報告書 (日本語) が不適切だと誤訳になる。★学習機能があると先走る可能性あり</p>	<p>他の薬剤名が記載される。 原文にない記載が自動的に追加された箇所 (表タイトル) が複数ある。 一部のフレーズの訳漏れがある。</p>

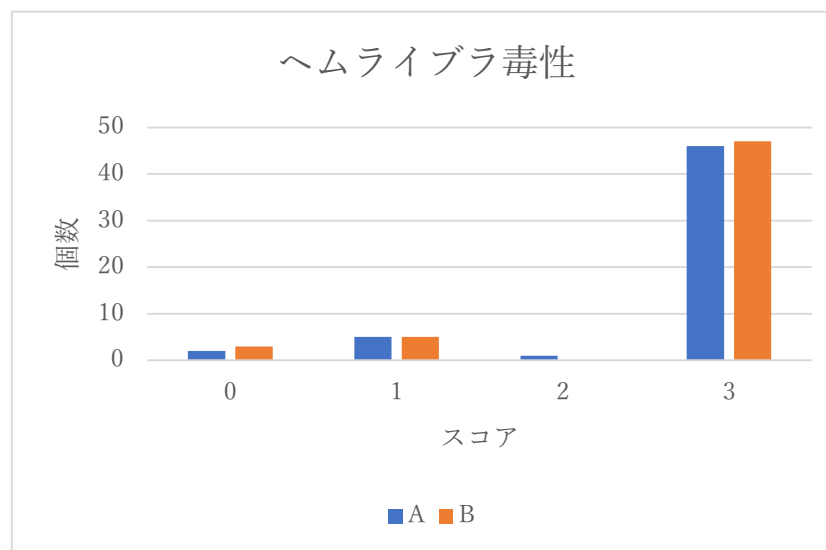
【図6】 スキリージ英訳文書品質評価スコア分布



【表9】 ヘムライブラ英訳文書の毒性分野評価まとめ

ツ ー ル	スコア (件数)				計	図表について	0 or 1 のおもな理由
	0	1	2	3			
A	2	5	1	46	54	【表・図】 なし。	毒性分野に頻出する用語の翻訳が不正確。 原文が難解で正しく翻訳されない。
B	3	5	0	47	55	【表・図】 なし。	原文が難解で正しく翻訳されない。

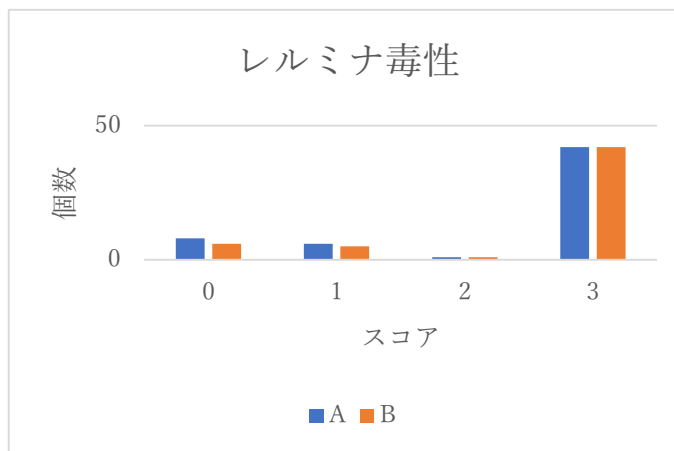
【図7】 ヘムライブラ英訳文書品質評価スコア分布



【表10】レルミナ英訳文書の毒性分野評価まとめ

ツ ー ル	スコア（個数）				計	図表について	0 or 1 のおもな理由
	0	1	2	3			
A	8	6	1	42	57	<p>表・図】表11 ★ALK・CKの増加 → ALK/CK increased。表12 ★腎臓の尿細管上皮の壊死、好塩基性尿細管 → Necrosis and basophilic tubular epithelium of kidney(誤訳) 腎臓の壊死と好塩基性尿管…となる。★核片貪食→karyophagic cell となっているがこれは核貪食、nuclear fragments phagocytosis ではない。★文末に「etc.」が入ってる。表14「諸組織」が(意味不明) →Various organizations。</p>	<p>原文が長文で複雑なため正確に翻訳されない。毒性分野で頻出する用語、表現の訳や使い分け(受胎能、生殖能、fertility, fecundity 等)が適切でない。</p>
B	6	5	1	42	53	<p>【表・図】表11★Administration Pathway → Dosing Route (以下同様)。★表15 親動物、母動物の用語統一必要。</p>	<p>1文に()が14か所等、難解な文。また、文中に()があると誤訳が多い。毒性分野で頻出する用語、表現の訳が適切でない。訳漏れ、構文の不正確さ。</p>

【図8】レルミナ英訳文書品質評価スコア分布



【表 1 1】 各翻訳産物の毒性分野で 2 あるいは 3 と評価されたセンテンス数及びその割合

審査報告書	イミフィンジ		スキリージ		へムライブラ		レルミナ		合計	
	A	B	A	B	A	B	A	B	A	B
翻訳ツール										
スコア 2 及び 3 のセンテンス数	39	38	17	19	47	47	42	42	145	146
評価対象全センテンス数	51	52	35	37	54	55	57	53	197	197
割合 (%)	76.5	73.1	48.6	51.4	87.0	85.5	73.7	79.2	73.6	74.1

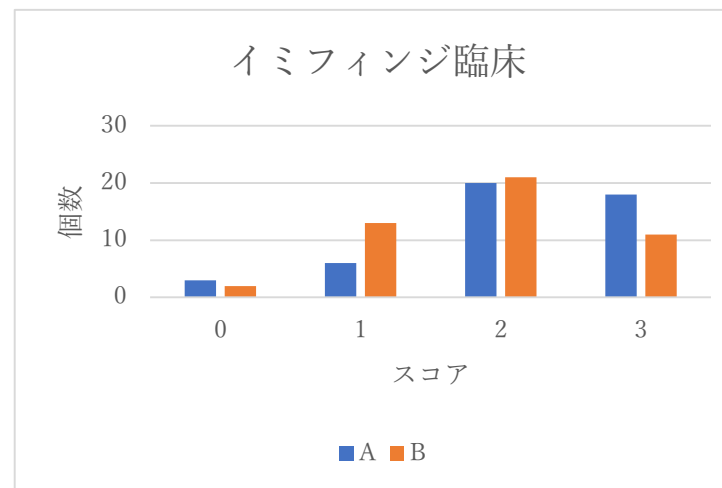
【表 1 2】 毒性分野の英訳不備内容の件数

対象文書 不備内容	イミフィンジ		スキリージ		へムライブラ		レルミナ	
	A	B	A	B	A	B	A	B
訳漏れ	7	3	1	3	0	0	1	2
原文の誤解釈	5	10	4	8	2	3	6	8
訳文の文法／構文	5	5	3	0	3	2	7	5
不適切用語	2	4	5	4	3	4	2	3
記号・数字	1	2	3	2	1	0	3	1
その他	0	0	5	6	0	0	0	0
合計	20	24	21	23	9	9	19	19

【表13】 イミフィンジ英訳文書の臨床分野評価まとめ

ツ ー ル	スコア（個数）				計	概要	0 or 1 のおもな理由
	0	1	2	3			
A	3	6	20	18	47	精度は良い。表については、数字を落とすことなく正確に作表できていた。 肯定/否定の逆転等の Critical な問題は検出されなかった。	原文に無い数字の挿入が2か所。 二重否定が訳せていない。 生体反応の表現が不適切。 訳漏れ、主語の誤り 等。
B	2	13	21	11	47	数字はミスなく出力。大きな訳抜け（一文全部抜ける、等）なし。冠詞や単数/複数形のミスが情報の正確さに影響。	フレーズをつなぐ語の不足。構文の誤解釈。「図1」が「表1」に誤訳。図の軸名、データ項目名が無い。死因羅列で訳抜けで対象数値がずれる。定冠詞や一部フレーズ等の細かい訳漏れが意味に影響する箇所が多数あり。

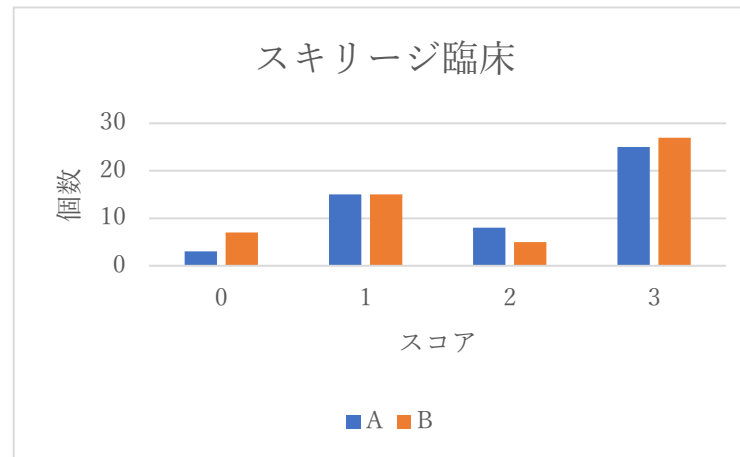
【図9】 イミフィンジ英訳文書臨床評価スコア分布



【表14】 スキリージ英訳文書の臨床分野評価まとめ

	スコア (個数)				計	概要	0 or 1 のおもな理由
	0	1	2	3			
A	3	15	8	25	51	<p>文章の途中に図表番号や脚注番号が入ると正確な訳出ができなくなる傾向。</p> <p>有害事象名は MedDRA 等と照合せず訳語として適・不適の観点で評価。</p>	<p>主語が誤り。数字の後等で訳漏れ。図表番号前後で文章が崩れる。逐語の直訳による不自然な訳。「上咽頭炎」「鼻咽頭炎」が同じ訳とされ区別できない。同一用語の訳が不安定。カッコ内の数値が不正確。無関係な薬剤名が出てくる。Table 番号が不自然に挿入。元文にない文が挿入される。「目的」を「Eyes Common」と誤訳。</p>
B	7	15	5	27	54	<p>元文が長いと訳文にも区切り入らず不自然になる傾向。</p> <p>有害事象名は MedDRA 等と照合せず訳語として適・不適の観点で評価。</p>	<p>「例数」表記に「N」と「subject」が混在等、訳が一定しない。</p> <p>表中の数字の誤訳。脚注での不要な用語の挿入。</p> <p>「リスク管理計画」の誤訳。</p>

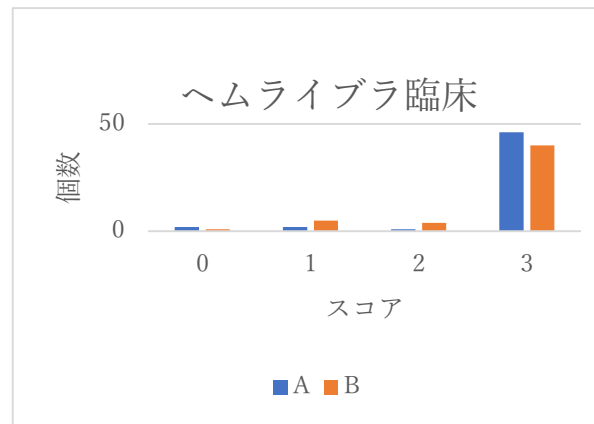
【図10】 スキリージ英訳文書臨床評価スコア分布



【表15】ヘムライブラ英訳文書の臨床分野評価まとめ

	スコア（個数）				計	概要	0 or 1 のおもな理由
	0	1	2	3			
A	2	2	1	46	51	<p>全体的に文章が自然で読みやすくポストエディットしやすい。専門用語を正しく選択している。</p> <p>同じ用語の訳にブレがある（製剤の訳など）。</p> <p>元文に多重の丸括弧があると誤訳に繋がる可能性。複数種類のカッコを用いる必要あり。</p> <p>0例を0 subjects と単数形になっていない。</p> <p>「表〇〇のとおりであった。」が「Table 〇〇Yes」と訳されている箇所が複数。</p>	<p>多重丸カッコ内の解析対象例数が不正確。</p> <p>「Table 19 に示す」を「were:Table 19Yes.」と訳。</p> <p>内容理解に影響する訳漏れ。1文訳漏れ。</p>
B	1	5	4	40	50	<p>同じ用語の訳にブレがある（製剤の訳など）。</p> <p>元文で多重の丸括弧があると誤訳に繋がる可能性。複数種類のカッコを用いる必要あり。</p> <p>0例を0 subjects と訳しているが、単数形にすべき。</p>	<p>多重丸カッコ内の解析対象例数が不正確。</p> <p>内容理解に影響する訳漏れ：単位の訳漏れ、「7件4例」の4例訳漏れ等</p>

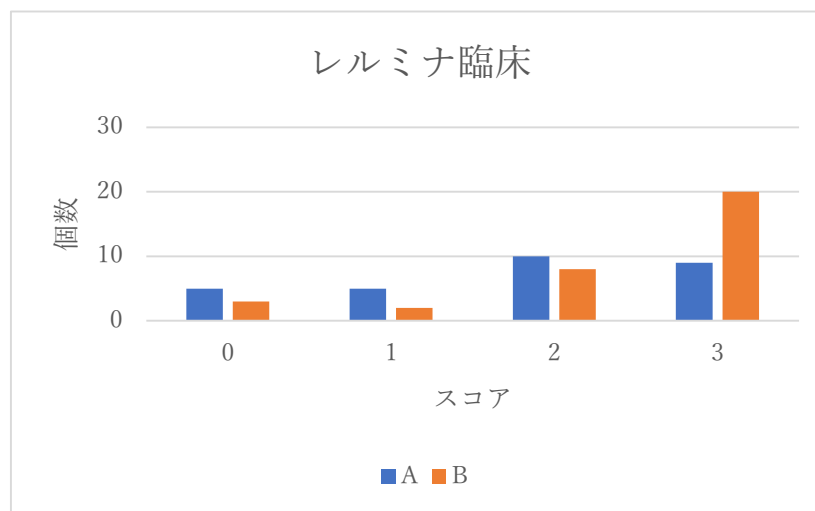
【図11】ヘムライブラ英訳文書臨床評価スコア分布



【表 1 6】レルミナ英訳文書の臨床分野評価まとめ

	スコア（個数）				計	自由記載（概要）	0 or 1 のおもな理由
	0	1	2	3			
A	5	5	10	9	29	元文が長文の場合、句読点が適切でなかったり複雑な構成となり、重大な誤訳や訳漏れ、構文の乱れが生じやすい。文の途中で引用先をカッコ付で記載すると、そこで文を分割して訳され前後のつながりがわかりにくくなる場合が複数ある。 「本薬」を「the this drug group」。	解析からの除外を試験からの除外と訳。上付文字が普通の大きさに注釈が伝わらない。 元文を不適切な個所で分割したため訳文の位置付けがわかりにくい箇所が複数ある。 元文が長く句読点が不適切なため重要な部分が訳漏れ。
B	3	2	8	20	33	「本薬」が「drug」と表記されているためわかりにくい箇所が複数ある。 元文が長文の場合は分割されるが、元文の論旨が複雑あるいは不明瞭な場合は誤訳が発生しやすい。	「本薬」を「drug」とし意味が通じにくい。 構文が複雑だと誤解釈や訳漏れが発生。分割されて訳の意味が混乱する。

【図 1 2】レルミナ英訳文書臨床評価スコア分布



【表 17】 各翻訳産物の臨床分野で 2 あるいは 3 と評価されたセンテンス数及びその割合

	イミフィンジ		スキリージ		ヘムライブラ		レルミナ		合計	
	A	B	A	B	A	B	A	B	A	B
スコア 2 及び 3 の個数	38	33	33	32	47	44	19	28	137	137
全センテンス数	47	47	51	54	51	50	29	33	178	184
割合 (%)	80.9	70.2	64.7	59.3	92.2	88.0	65.5	84.8	77.0	74.5

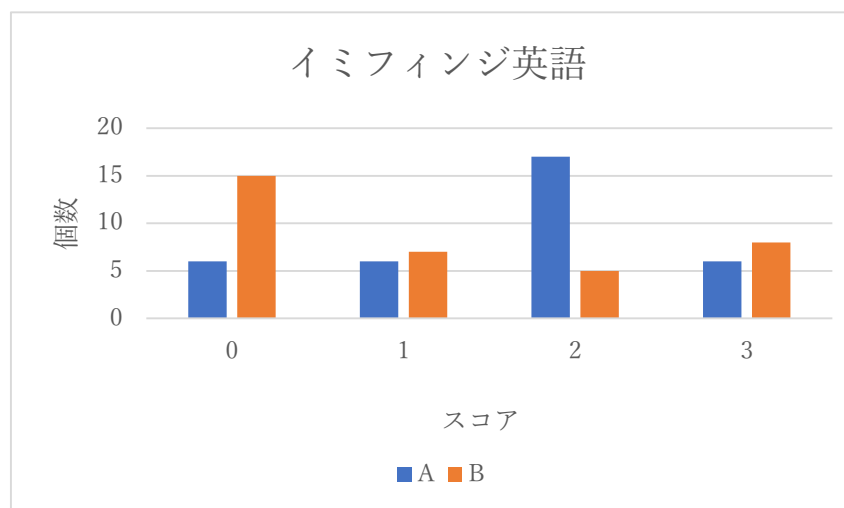
【表 18】 臨床分野の英訳不備内容の件数

対象文書 不備内容	イミフィンジ		スキリージ		ヘムライブラ		レルミナ	
	A	B	A	B	A	B	A	B
訳漏れ	6	9	9	4	2	2	4	0
原文の誤解釈	1	2	4	1	1	1	8	9
訳文の文法／構文	17	27	12	10	3	0	9	2
不適切用語	2	10	2	9	0	11	0	2
記号・数字	5	3	0	1	0	0	1	0
その他	7	1	7	16	0	0	0	0
合計	38	52	34	41	6	14	22	13

【表19】イミフィンジ英訳文書の英語の正確性の視点での評価まとめ

	0	1	2	3	計	0 or 1 のおもな理由
A	6	6	17	6	35	「① 0.1 又は②1mg/kg」を「10.1 or 201 mg/kg」等、項目番号と数値をすべて数値として訳した箇所が複数。 上付、下付文字が普通の大きさと翻訳され混乱。 やや長文を不適切な個所で分割したことで原文構成の誤解釈や訳漏れ発生。 Imifinzi を Imifindi とスペルミス。
B	15	7	5	8	35	「本薬の用法・用量について」を「Dosage and administration of ramucirumab」と関係ない薬剤名が挿入されている。 Imifinzi、durvalmab を Imifindi、dulvalumab とスペルミス。 長めの文で訳漏れやご解釈が散見されるが、長文でなくても1フレーズの訳漏れが複数ある。 統合解析を「combined analysis」と訳。Intedrated anasysis

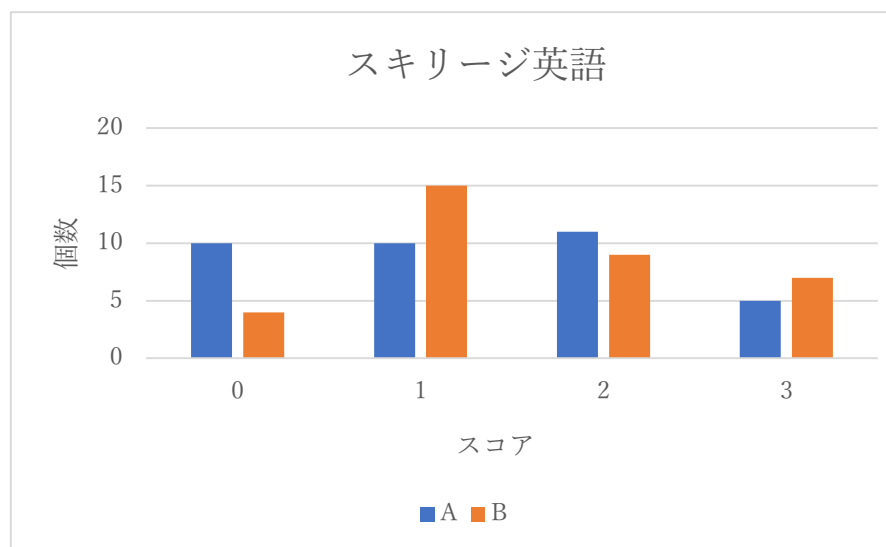
【図13】イミフィンジ英訳文書の英語の正確性の視点でのスコア分布



【表 2 0】 スキリージ英訳文書の英語の正確性の視点での評価まとめ

	0	1	2	3	計	0 or 1 のおもな理由
A	10	10	11	5	36	<p>カッコ内の記載の訳がカッコ外にある。</p> <p>「Table○ Yes」が頻出。</p> <p>脚注番号の前後で文を分割しているが文法的に不完全な訳。数値とカッコが続く文で訳漏れ。</p> <p>複数投与群の説明で用量と症例数の数値、群の名称の番号が入れ替わる等、数字が多い文で入れ替わる箇所が多数ある。</p> <p>グラフ画像の日本語がそのまま。原文にない数値等の挿入。</p>
B	4	15	9	7	35	<p>マスキングとカッコが重なる記載、カッコや数値を含む長文で構文の乱れや訳漏れする場が複数あり。</p> <p>薬物動態、血中薬物濃度、皮下投与、抗体価の訳が不適切。同一表中含め訳語のブレ。</p> <p>脚注番号が別の文に移動する例が複数あり。</p>

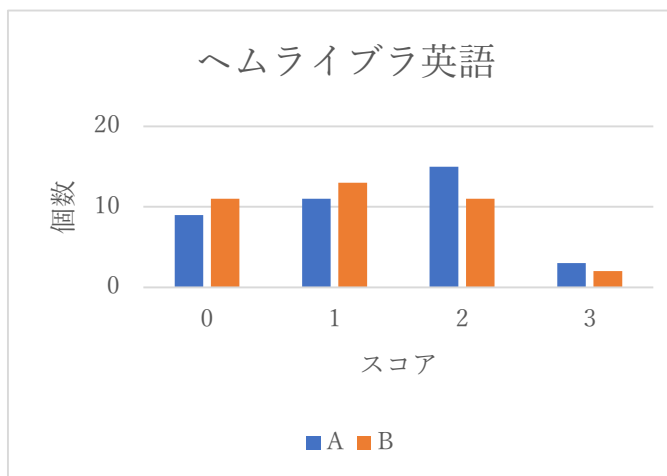
【図 1 4】 スキリージ英訳文書の英語の正確性の視点でのスコア分布



【表2 1】ヘムライブラ英訳文書の英語の正確性の視点での評価まとめ

	0	1	2	3	計	0 or 1 のおもな理由
A	9	11	15	3	38	Hemlibura を Hemribula とスペルミス。[[目 次]を[Eyes] Next}と訳。 「同〇mg、同〇mg、同〇mg、…」で1か所「mg」の訳漏れ。「本質」を「Reference」、「構造」を「Design」等、審査報告表紙等の定型記載の訳が不適切。糖鎖構造の「Glc」「GlcNAc」が訳漏れ。「別記の通り。」を「Appended listYes.」。 マスキングが多い文では訳の構文に乱れがありマスキング箇所の位置が不適切。長文で訳漏れ。
	11	13	11	2	37	「ヘムライブラ」を「Heme Library」、「本質」を「Essential」、「構造」を「Preparation」と不適切な用語に翻訳。 記号、数値が多数含まれる糖鎖構造の説明で、訳漏れやハイフンの漏れ等が多数発生。 希少疾病用医薬品指定通知の記載が不適切。審査部名の記載が不適切。目次が日本語のまま。 「1週間の間隔で4回皮下投与」を「4 times every week」 原文が複雑な構文の場合に訳漏れ。

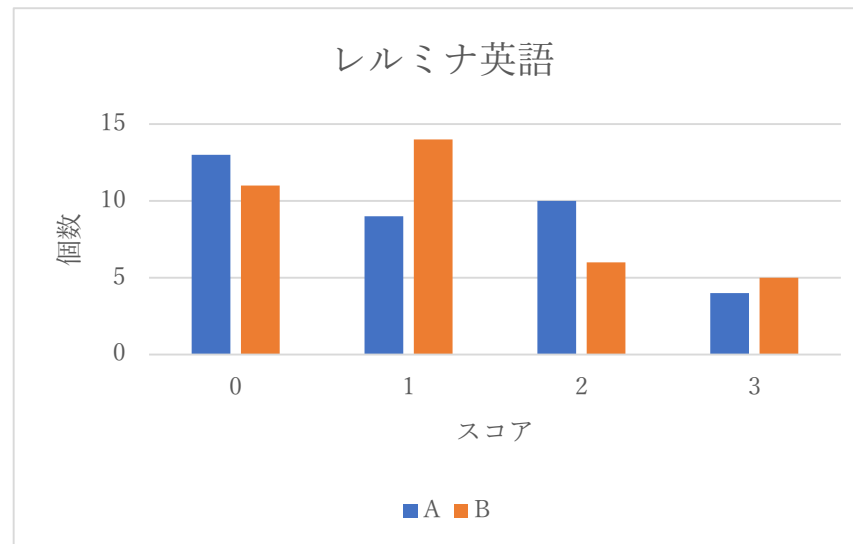
【図1 5】ヘムライブラ英訳文書の英語の正確性の視点でのスコア分布



【表 2 2】レルミナ英訳文書の英語の正確性の視点での評価まとめ

	0	1	2	3	計	0 or 1 のおもな理由
A	13	9	10	4	36	原文の構成が不適切な構成のため訳もわかりにくい。長文の場合に原文の誤解釈が発生したり、不適切な個所で分割して誤訳につながるケースが多数。 数値を含むカッコが多数あると数値の入れ替わりや訳漏れが複数発生する事例が複数あった。略語表の整備が必要。
B	11	14	6	5	35	原文が不適切な構成のため、原文の誤解釈が生じて訳の意味が変わる、訳漏れが発生する、わかりにくい等の箇所が多数みられる。臨床症状等の訳が不適切。

【図 1 6】レルミナ英訳文書の英語の正確性の視点でのスコア分布



【表 2 3】 各翻訳産物の英語の正確性の視点で 2 あるいは 3 と評価されたセンテンス数及びその割合

	イミフィンジ		スキリージ		へムライブラ		レルミナ		合計	
	A	B	A	B	A	B	A	B	A	B
スコア 2 及び 3 の個数	25	13	16	16	18	13	14	11	73	53
全センテンス数	36	35	36	35	38	37	35	35	145	142
割合 (%)	69.4	37.1	44.4	45.7	47.4	35.1	40.0	31.4	50.3	37.3

【表 2 4】 英語の正確性の視点での評価における英訳不備内容の件数

対象文書 不備内容	イミフィンジ		スキリージ		へムライブラ		レルミナ	
	A	B	A	B	A	B	A	B
訳漏れ	3	9	4	6	4	5	6	9
原文の誤解釈	5	12	5	6	6	10	14	18
訳文の文法／構文	1	4	9	7	1	1	4	4
不適切用語	11	11	7	12	5	3	12	14
記号・数字	6	4	10	8	3	2	9	7
その他	11	6	20	20	25	25	6	6
合計	37	46	55	59	44	46	51	58

令和元年度厚生労働行政推進調査事業費補助金（厚生労働科学特別研究事業）
（分担）研究報告書

ADME分野等検討

研究分担者 永井 尚美 武蔵野大学薬学部・教授

医薬品等の開発・流通に関わる共通課題を解決してグローバルヘルスに貢献するために、規制当局間の緊密な連携は不可欠であり、日本から必要な情報発信を適時に行う必要がある。本研究では、ADME 分野の医薬品評価情報作成における AI 自動翻訳の有用性を評価し、審査報告書の英語版作成における自動翻訳産物の有用性と限界、活用時の留意点について検討を行った。

新医薬品の審査報告書の ADME 分野について、AI 自動翻訳ツールによる英語版の翻訳の質を、専門性の観点から評価し、原文の解釈及び翻訳版での文意伝達に関わる翻訳ツールの特徴、翻訳産物の利用にあたっての留意点を整理した。自動翻訳ツールによる特徴的な翻訳不備の傾向を理解したうえで、文意伝達の点で重大な問題が生じる部分について、専門分野及び言語の正確性の観点からのポストエディットにより翻訳作業の迅速化に資すると考えられた。

A．研究目的

医薬品等の開発・流通のグローバル化と製品の多様化により、規制上の共通課題を解決してグローバルヘルスに貢献するために、規制当局間の緊密な連携は不可欠なものとなっている。海外規制当局に対して、必要な情報発信を適時に行うためには、専門的な内容を含む報告書等を正確かつ迅速に翻訳する必要があるが、このための人的リソースの確保には限界がある。この問題の改善策として、AI自動翻訳という新たなツール導入が考えられる。

ADME分野の規制に関しても、臨床開発の早期段階からの相談、ICHでの複合領域（M）ガイドラインの作成、申請電子データ対応を含めた承認審査等、広範囲の案件について、海外規制当局への情報発信や緊密な連携が重要になっている。必要なADME医薬品評価情報を迅速に翻訳して発信する機会は、今後益々増大すると考えられる。

本研究では、ADME分野における自動翻訳ツールの評価を行うとともに、他の専門分野との横断的な議論も踏まえ、医薬品評価情報作成における自動翻訳の有用性と限界、利用にあたっての留意点を評価することを目的とした。

B．研究方法

1. 自動翻訳ツールの選定

本研究で使用する自動翻訳ツールとして、医学・薬学の専門用語に対応可能なAI等を活用したT-400（Translation For Onsha Only：株式会社ロゼッタ）及びCOTOHA（株式会社みらい翻訳/NTTコミュニケーションズ/株式会社翻訳センター）を選定した。なお、本研究においては2つの自動翻訳ツール間の性能比較が目的ではないため、自動翻訳ツール名を【A】あるいは【B】と記載する。

2. 対象品目の選定と学習

医薬品医療機器総合機構（PMDA）では、新医薬品の審査報告書の一部を英訳してホームページに掲載している。

（<http://www.pmda.go.jp/english/review-services/reviews/approved-information/drugs/0001.html>）

承認医薬品の品質特性、非臨床評価手法、疾患領域等の網羅性に配慮しつつ、一定数の審査報告書の英訳版を自動翻訳ツールで学習させた。自動翻訳ツールで学習させていない審査報告書を複数の疾患領域から選定して評価対象とし、自動翻訳ツールにより英訳版を作成した。悪性腫瘍、産科婦人科、炎

症性疾患及び血液疾患領域から、以下に示す化学合成医薬品又はバイオ医薬品(経口固形製剤又は注射用製剤)計4品目の審査報告書を評価対象とした。

- イミフィンジ点滴静注 120mg、同点滴静注 500mg (以下、イミフィンジ)
(https://www.pmda.go.jp/drugs/2018/P20180727001/670227000_23000AMX00485_A100_1.pdf) 抗悪性腫瘍剤
- スキリージ皮下注 75mg シリンジ 0.83mL (以下、スキリージ)
(https://www.pmda.go.jp/drugs/2019/P20190329001/112130000_23100AMX00299_A100_1.pdf) 炎症性疾患治療薬
- ヘムライブラ皮下注 30mg、同皮下注 60mg、同皮下注 90mg、同皮下注 105mg、同皮下 150mg (以下、ヘムライブラ)
(https://www.pmda.go.jp/drugs/2018/P20180329003/450045000_23000AMX00450_A100_1.pdf) 血液疾患治療薬
- レルミナ錠 40mg (以下、レルミナ)
(https://www.pmda.go.jp/drugs/2019/P20190124001/400256000_23100AMX00010_A100_1.pdf) 婦人科疾患治療薬

3. 自動翻訳の有用性評価

自動翻訳版の翻訳精度及び医薬品評価情報作成における自動翻訳ツールの有用性をADME分野の医薬品評価の観点から検討した。

ADME分野の審査報告書は、非臨床及び臨床に分けて記載される。ADME評価情報として、医薬品開発で実施される典型的な薬物動態試験、臨床薬理試験、生体試料分析法、薬物動態等の解析方法、審査において一般的に議論される重要事項が網羅されるよう、下記の内容を選択した。そのうえで、非臨床及び臨床両パートを対象品目に振り分け、誤訳が生じやすいと考えられる箇所を中心に、専門用語、項目及び図表を含む評価対象箇所を抽出した(表1)。

1. 非臨床薬物動態

分析法、組織分布、代謝及び排泄、反復投与試験とトキシコキネティクス

2. 臨床薬理

分析法と製剤設計、代謝に関わる酵素分子種、バイオアベイラビリティ試験、第 相試験、国際共同試験、母集団薬物動態解析、曝露-応答解析

3. 審査内容

臓器障害(腎機能及び肝機能の低下)が薬物動態に及ぼす影響、薬物動態における国内外差、用法・用量設定における薬物動態・臨床薬理の観点からの考察、薬物動態学的相互作用

自動翻訳ツールを用いて翻訳された文書の評価対象箇所について、1文ごとの評価を行った。

ADME分野については、当該専門性を有する評価者1名が、4品目の審査報告書の【A】及び【B】による翻訳産物を評価した。

翻訳精度の評価は、4段階のスコア「0:重大な不備あり(全く意味が通らない)」、「1:修正が必要な不備あり(誤解を招く、或いは意味が伝わらない)」、「2:修正不要の不備あり(意味は十分伝わる)」及び「3:問題なし」を設定して採点形式で実施した。スコア2及び3とした割合に基づき、専門性の観点からの審査報告書の解釈及び翻訳版での文意伝達に関わる総合的な判断を行った。翻訳産物における意味伝達のうえでの重大な不備、誤解を招く或いは原文が意図する内容が十分伝わらず修正を要するとして、スコア0又は1と判断した背景情報を整理した。

これら判断に基づき、自動翻訳の有用性や限界、自動翻訳を用いる際の留意点について、他の分野の専門性を有する評価者との意見交換も踏まえて有用性評価を行った。

なお、自動翻訳ツール及び対象品目の選定、翻訳精度の評価方法は全分野で共通であり、詳細は総括研究報告書及び品質・毒性・臨床分野等検討を参照いただきたい。

(倫理面への配慮)

自動翻訳ツールの採用の経緯ならびに評価方法及び結果を本報告書で公表し透明性を確保することで利益相反が無いことを示す。

また本研究は、再生医療等の安全性の確保等に関する法律、人を対象とする医学系研究に関する倫理指針、ヒトゲノム・遺伝子解析研究に関する倫理指針、動物実験等の実施に関する基本指針、省令GCP、臨床研究法の対象に該当しない。

C. 研究結果

4つの品目の審査報告書のADME分野について自動翻訳ツール【A】及び【B】により翻訳した産物から、ADME評価情報として、図表計13を含む報告書本文及び項目タイトル計188箇所のスコア的情報を図1～4に示す。また、ADME評価情報のうち、審査報告書の解釈及び翻訳版での文意伝達に関して、専門的立場からは修正しなくても意味が通ると評価されたスコア2及び3と評価された箇所及びその割合を表2にまとめた。スコア2又は3と判断された箇所の割合は、67～94%の範囲であり、スキリージで60%後半、ヘムライブラとイミフィンジで90%前後、レルミナで80%前後であった。全188箇所における平均は、【A】及び【B】の翻訳産物で83%及び85%であった。同一品目においては、翻訳ツール間の差は最大で7%程度であり顕著な差は認められなかったが、品目間では67%から94%の範囲となった。

全体を通じて、主述の関係、修飾語と被修飾語との関係が明瞭な原文については、翻訳産物においても比較的抵抗なく文脈を追うことができ、一文全てが訳抜けといった文意伝達のうえでの重大な問題は認められなかった。また、項目タイトルや文中の専門用語の選択も概ね正確であり、非臨床及び臨床薬物動態情報、並びに薬物動態等解析結果を示す典型的な図表の形崩れも認められなかった。

翻訳不備の内容ごとの件数を表3に示した。ヘムライブラを除く3品目で「原文の誤解釈」の件数が多く、原文が難解な長文であることが原因と考えられた。新又は旧製剤の表現として、new・old・current productの混在、群の表現として、armとgroup、その他に上腕部、大腿部、出生児等、今回検討した評価対象箇所内で認められた同一用語の訳ぶれについては「その他」に分類したうえで、文意伝達のうえで問題ない場合はスコア2と判断し

た。また、審査報告書で共通に用いられる「本薬」及び「本剤」の翻訳産物として、特定の固有名詞(医薬品名)を挙げている箇所は「不適切用語」として個々に特定したが、this drugやdrug等の表現が評価対象範囲を通じて出現する点は、全体に係る事項として整理した。

スコア0及び1の翻訳不備と判断した事由として、記号・数字の不備、不適切な用語選択、訳漏れ・訳抜け、訳ぶれ、原文の誤解釈、文法・構文の不備が認められた。各々の事由の背景情報を以下に記す。

1. 「記号・数字の不備」「不適切な用語選択」
 - 有効成分及び製剤の表現として、全ての審査報告書で共通に用いられる「本薬」及び「本剤」に関わる用語選択として、他の医薬品名(固有名詞)の唐突な出現が認められた。
 - 丸括弧を数値として解釈したり、自動的にゼロを追加して異なる数値に変換したりする等の不備が認められた。
 - 上付、下付、中点等への対応に一部不備が認められた。
 - 「absorption」を「resorption」、「isozyme(又はisoform)」を「isomer」等、一部のADME関連領域の用語選択が不適切であった。
2. 「訳漏れ・訳抜け」「訳ぶれ」
 - 内容理解に影響する数字や単位における不規則な訳漏れ・訳抜けが認められた。
 - 内容理解に影響する重要な名詞・形容詞の訳漏れ・訳抜けが認められた。
 - 新旧製剤の表現において、new・old・currentの対応が正確でない箇所が認められた。
3. 「原文の誤解釈」「文法・構文の不備」
 - 測定項目と対応すべきデータ、薬物と対応すべき用量情報の不一致が認められた。
 - 文中又は文末に図表タイトルや図表・脚注番号が不自然・不規則に挿入又は付記される箇所が認められた。
 - 本文中への多重の括弧や図表・脚注番号の挿入により、翻訳の正確性への影響が認められた。
 - 原文が長文の場合には、括弧や記号が存在する前後において、自動的に一文を複数に分割して

翻訳する傾向が認められ、このような箇所では、翻訳文の文法・構文が不自然になり、文意伝達に重大な影響を与える場合が認められた。

- 原文が長文の場合には、修飾語と被修飾語との関係、主述の関係が適切に翻訳されていない箇所が認められた。
- 健康被験者とすべき箇所を患者、表中の項目において、非臨床薬物動態データであるにも関わらずヒトのデータとして取り扱われている箇所があった。
- 日本語特有の省略、冠詞や単複数の別といった日本語にはない文法上の取り扱いへの対応が必ずしも十分ではなく、翻訳の際に評価情報の正確な伝達に影響する場合が認められた。

以上より、特に複雑な原文構造或いは日本語特融の省略等により論旨が不明瞭な長文については、翻訳の際に文章が自動的に複数に分割され、訳漏れ・訳抜け、不自然な構文、原文の誤解釈による重大な誤訳が発生する可能性があると考えられた。

D . 考察

1 . ADME評価情報について

評価の対象とした疾患は、悪性腫瘍、産科婦人科、炎症性疾患及び血液疾患の4領域であったが、PMDAの審査部・チーム横断的な疾患領域の選択がなされた。対象医薬品の有効成分は化学合成品又はバイオ医薬品、投与経路は経口又は静脈内、剤形は固形又は注射用製剤であり、ADME審査において特に着目すべき重要な事項の大部分を網羅していると考えられた。また、審査報告書から選定した専門用語、項目及び図表を含むADME評価対象箇所には、医薬品開発過程で実施される非臨床及び臨床領域の典型的な薬物動態及び臨床薬理試験、生体試料分析法、薬物動態等解析方法を含み、ADME審査における主要な議論内容を概ね網羅すると考えられた。

今回の評価対象箇所は、一部の疾患領域におけるADME分野報告書の部分抽出であり、他の疾患領域や審査報告書全体に一般化するには限界があると考えられるが、ADME分野に関わる自動翻訳精度

の有用性評価と課題抽出との観点からは適切な選定だったと考えられた。

2 . 自動翻訳の評価手法と有用性評価について

本研究で採用した評価手法は、翻訳精度を4段階とし採点方式による評価を行ったうえで、スコア2と3を統合した割合に基づき、自動翻訳ツールによる審査報告書の解釈及び翻訳産物の文意伝達に関わる総合的な判断を行った。また、スコア0及び1と判断した個別の背景情報を分析することにより、個々の翻訳不備の事由・背景情報を精査した。ADME分野については、当該領域の専門性及び医薬品評価の経験を有する1名の評価者が上述した評価手法により検討を行った。

ADME領域の専門用語の選択において、一部の用語の翻訳が不正確であったり、訳ぶれが生じていたり、また、意味不明な用語の突発的な挿入が数か所認められたものの、ADME評価情報を読み進めるにあたって大きな支障を感じることはなかった。また、構文や文脈が明瞭な日本語原文については、翻訳産物に基づく文意理解のうえで特段の支障を感じることはなかった。この点については、今回検討した翻訳ツール間で概ね同様であった。

スコア2と3を統合した割合は、同一品目においては、翻訳ツール間の差は最大で7%程度であり顕著な差は認められなかったが、品目間では67%から94%の範囲となった。この背景として、今回選定した非臨床薬物動態、臨床薬理及び審査内容が品目ごとに異なっていること、それに伴い評価情報における図表の有無や抽出箇所数が異なることが考えられた。スキリージのスコア2及び3の割合が他品目と比較して低値であった。本品目の評価情報として抽出した非臨床薬物動態の箇所は反復投与試験(トキシコキネティクス)であり、表を含めて総評価情報に占める割合が高かった。トキシコキネティクスは、毒性試験の情報として得られ、用語や評価の視点は毒性領域と関わりがある。また、臨床薬物動態の評価対象としては、国際共同試験や曝露-応答解析及び用法・用量設定に関わる内容であった。スキリージのスコア2及び3の割合は、毒性及び臨床分野の評価と対応すると考えられた。また、同一品

目においては、翻訳ツール間に顕著な差は認めなかったが、スコア2及び3の割合に最大で7%程度の差を認めたレルミナは、臨床分野の評価と同様の傾向であった。

原文が長文であり、特に日本語特有の省略又は執筆者の個性による複雑な構文については、両翻訳ツールともに自動的に一文を複数に分割して翻訳する傾向が認められ、その際に訳漏れ・訳抜け、主述又は修飾・被修飾の関係の変化が生じる場合が多いと考えられた。また、文中又は文末に図表タイトルや図表・脚注番号が不自然に挿入又は付記される場合、「本薬」や「本剤」がdrug等の一般的な名称で訳出される場合が認められた。これら系統的な訳出傾向に加え、「本薬」や「本剤」が特定の固有名詞(医薬品名)として突発的に訳出される箇所も認められた。これらの点は、データ及びPMDAの評価情報を正確に伝達するとの観点からは重大な問題であると考えられた。

今回選定した抽出箇所が品目ごとに異なることに加え、ADME分野のみが非臨床及び臨床の両領域に関わることから、ADME以外の専門分野の医薬品評価や疾患領域の情報が翻訳精度に影響することが示唆された。また、今回の検討は単一の評価者によるものであったが、関連する他の専門分野における評価と齟齬はないと考えられた。以上より、今回用いた評価手法により、専門的な観点からは、自動翻訳の有用性と限界について、客観的に評価・整理することが出来たと考えられた。

3. 自動翻訳を用いる際の留意点と課題

審査報告書において共通に使用される「本薬」及び「本剤」の翻訳が適切に対応されていなかった。これら用語が審査報告書中に出現する頻度を踏まえると早急な改善が必要である。審査報告書全体に関係する「本薬」及び「本剤」については、汎用性のある一定のルール作成が必要と考えられた。

ADME分野の一部の専門用語の適切性及び訳ぶれは、翻訳ツール側の性能改善、同一疾患領域の医薬品、評価方法や品目特性が類似の医薬品についての学習と使用経験の蓄積により解決可能であると考えられた。専門用語については、標準的なルー

ル或いは翻訳辞書の整備が望まれる。また、ADME以外の専門分野の医薬品評価や疾患領域の学習経験と情報の蓄積はADME分野の翻訳精度向上にも資すると考えられた。

他分野同様にADME分野においても、本文への図表タイトル及び図表タイトル末尾への番号・記号の不規則・不自然な挿入・付記、特殊記号等の取り扱いが不適切な場合が認められた。これら技術的問題については、翻訳ツール側の性能改善により解決可能であると考えられた。

複雑な構造を有する文章及び言語体系に基づく日本語特有の省略等に対しては、現時点では、専門的なポストエディットが必須であると考えられた。

以上を踏まえ、PMDAが作成する審査報告書のADME分野について、翻訳実務への自動翻訳ツールの活用は可能であると考えられたが、今回の検討で明らかになった医薬品評価情報作成における自動翻訳の特徴を理解したうえで、集中的に人的支援を行う必要があると考えられた。また、PMDAにおける実務も踏まえ、審査報告書において、括弧による記載を避けたり、長文は複数に分割したりする等の対応、評価結果の表現の更なる標準化が期待された。

E. 結論

本研究では、ADME分野の医薬品評価情報作成における自動翻訳ツールの有用性と限界を評価し、利用にあたっての留意点を整理した。翻訳業務の質の保証のためには、現時点では、自動翻訳ツールの特徴と利用時の留意点を理解して適時適所なポストエディットを行う必要があると考えられた。今後、翻訳ツール側のより一層の性能改善及び人的支援との協働作業を整理することにより、AI自動翻訳は、正確かつ迅速な海外への情報発信に貢献することが期待された。

F. 研究発表

1. 論文発表

本研究に関連した論文発表はない。

2. 学会発表

本研究に関連した学会発表はない。

G. 知的財産権の出願・登録状況

(予定を含む。)

1. 特許取得

なし

2. 実用新案登録

なし

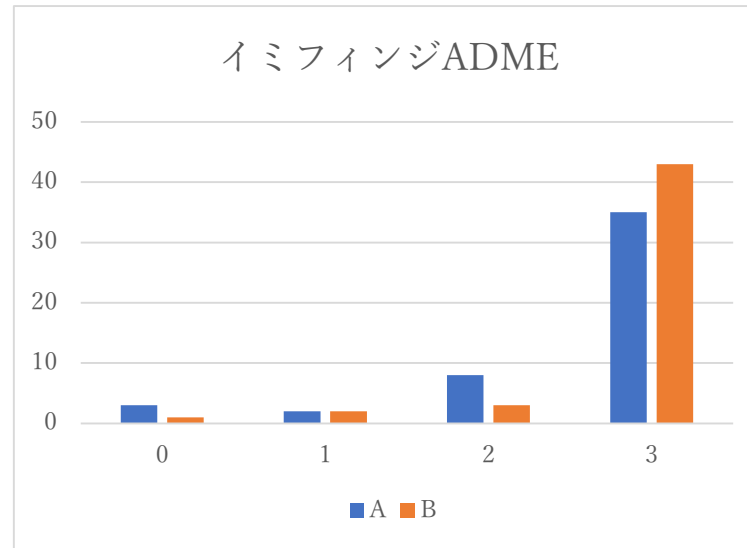
3. その他

なし

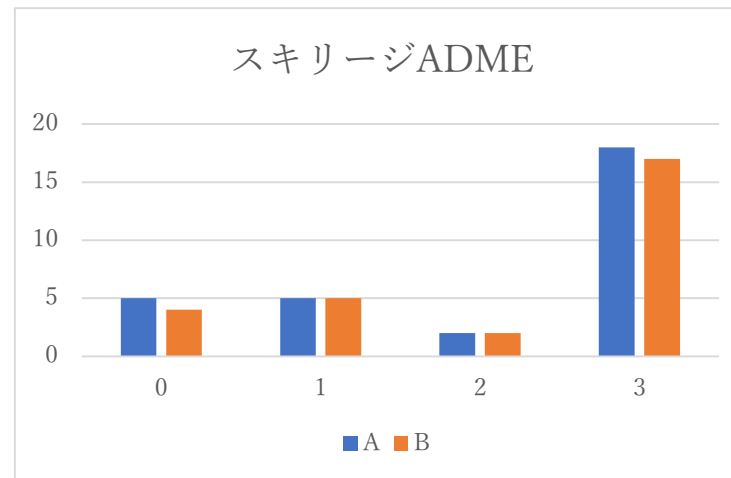
【表 1】 評価の対象とした文書及び箇所

文書	非臨床薬物動態	臨床薬物動態・臨床薬理	図表数
イミフィンジ 点滴静注	4.1 分析法～4.4 代謝及び排泄	6.1.1 分析法 6.1.2 開発過程における原薬及び製剤の製造工程の変更 6.2.6 腎機能及び肝機能の低下が本薬の PK に及ぼす影響、 6.R 機構における審査の概略、6.R.1 本薬の PK における国内外差	
スキリージ 皮下注	4.1.2 反復投与試験（トキシコキネティクス）	6.2.2.2 活動性 PsA 患者を対象とした国際共同試験 6.2.6.3 曝露量-応答解析 6.R 機構における審査の概略、6.R.1 国内臨床試験における用法・用量の設定について	表：4 図：1
ヘムライブラ 皮下注	4 非臨床薬物動態に関する資料及び機構における審査の概略	6.1.1 バイオアベイラビリティ試験 6.2 臨床薬理、6.2.1 国内第 I 相試験 6.R.1ADA について	表：6
レルミナ錠	4.2.1 組織分布（最初の段落） 4.R 機構における審査の概略、4.R.1 組織分布	6.2.1.2.2 本薬の代謝に関与する CYP 分子種の同定 6.2.4 母集団薬物動態解析 6.R.2 エリスロマイシンとの薬物動態学的相互作用について	表：2

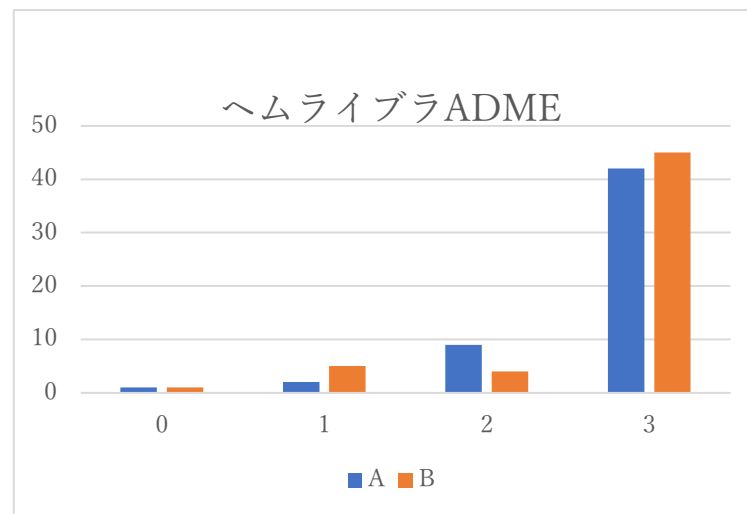
【図1】 イミフィンジ英訳文書 ADME 評価スコア分布



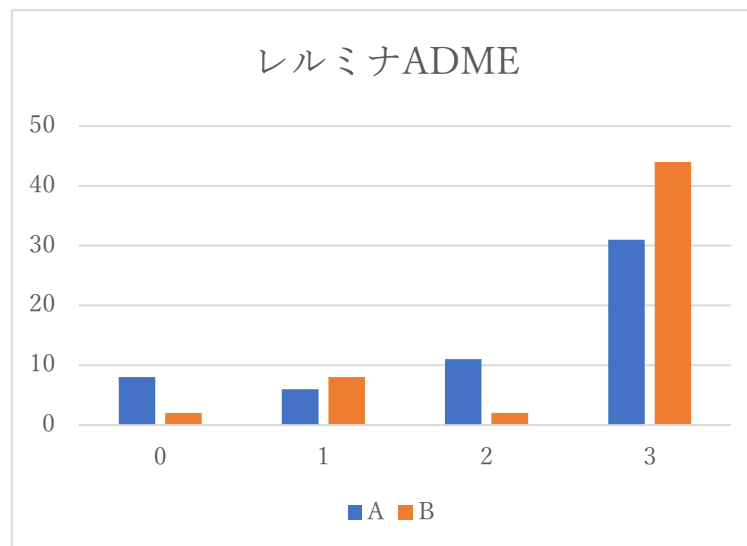
【図2】 スキリージ英訳文書 ADME 評価スコア分布



【図3】 ヘムライブラ英訳文書 ADME 評価スコア分布



【図4】 レルミナ英訳文書 ADME 評価スコア分布



【表2】各翻訳産物の ADME 分野で 2 あるいは 3 と評価された評価箇所の数及びその割合

	イミフィンジ		スキリージ		へムライブラ		レルミナ		合計	
	A	B	A	B	A	B	A	B	A	B
スコア 2 及び 3 の個数	43	46	20	19	51	49	42	46	156	160
全評価箇所数	48	49	30	28	54	55	56	56	188	188
割合 (%)	89.6	93.9	66.7	67.9	94.4	89.1	75.0	82.1	83.0	85.1

【表3】ADME 分野の英訳不備内容の件数

対象文書 不備内容	イミフィンジ		スキリージ		へムライブラ		レルミナ	
	A	B	A	B	A	B	A	B
訳漏れ	1	0	1	1	0	4	4	1
原文の誤解釈	5	3	5	6	0	1	8	3
訳文の文法／構文	2	0	8	6	2	0	2	3
不適切用語	0	0	1	0	3	1	4	5
記号・数字	6	0	4	4	9	4	8	1
その他	6	3	0	0	4	2	0	1
合計	20	6	19	17	18	12	26	14

令和元年度厚生労働行政推進調査事業費補助金（厚生労働科学特別研究事業）
（分担）研究報告書

薬理分野に関する検討

研究分担者 山田 博章 横浜薬科大学臨床薬学科・教授

研究要旨

新医薬品開発のグローバル化が進む中、日本から審査の情報をいち早く世界に発信する必要性が増してきている。今回、PMDAの新医薬品の審査報告書の英文への翻訳を効率化、高速化する手段として、最近、進歩が著しいAI翻訳システムの実用性を評価し、利用する場合の留意点についても検討した。

AI翻訳システムによる翻訳英文の約8割は、AI翻訳を利用したことを但し書きで断っておけば、情報伝達の目的において、許容可能な英文であった。残りの約2割の修正には、人の手による翻訳文全文の詳細なチェックが必須であるが、ある程度の対策を施すことによって、修正作業は省力化可能と思われる。従来の人の手によってゼロから翻訳していた工程をAI翻訳システムによる翻訳文のチェック作業に置き換えれば、英文化作業のかなりの効率化、高速化が可能になると考えられるが、データの解釈や固有名詞等については、審査担当者や審査経験者による最終チェックも必要と考えられる。また、実際にAI翻訳システムを導入する際には、審査報告書の日本語文章の特徴と照らし合わせて、親和性の高いシステムを検討する必要がある。

A. 研究目的

新医薬品の開発はグローバル化が進んでおり、国際共同治験による世界同時開発に日本が参加することも普通となってきている。そのような状況下において、PMDAによる新医薬品の承認審査や安全対策の情報をいち早く世界に発信する重要性が増してきているが、日本における審査報告書等の多くの情報は日本語で作成されているため、英文に翻訳するというステップが必要となってくる。最近では世界に先駆けて日本で最初に承認される新医薬品も増えつつあり、承認と同時に英文の審査報告書を世界に公開できることが望まれる。

現在、PMDAにおいて、主要な医薬品安全性情報に関しては、日本語とほぼ同日に英語でPMDAホームページに掲載されているが、新医薬品の審査報告書に関しては、各国にも影響を及ぼす可能性がある品目に優先順位をつけて英訳しており、承認してから公表までにかかなりの時間を要し、平成30年度の公表実績は37品目に留まる（平成30

事業年度業務報告, <https://www.pmda.go.jp/files/000232175.pdf> (2020.4.14 確認)。

一方において、コンピュータの性能は過去に比較して飛躍的に向上しており、ニューラルネットワーク技術を応用した深層学習による人工知能(AI)翻訳システムが商用化されている。これらのAI翻訳システムは従来の統計的機械翻訳と比べて、機械的な直訳ではなく、人が翻訳したものに近い精度の翻訳が得られるとされている(K. Go, *Mathesis Universalis*, 20, 27-41, 2019)。

今回、これらのAI翻訳システムを利用することによって、PMDAの審査報告書の英訳作業が効率化、高速化できることが期待されるため、PMDAの新医薬品の審査報告書の中から薬理領域の部分を2種類のAI翻訳システムを用いて翻訳された英文を対象に、AI翻訳の実用性と問題点を評価し、AI翻訳システムを利用する場合の留意点についても考察した。

B. 研究方法

PMDA のホームページで公開されているイミフィンジ点滴静注 120mg、同点滴静注 500mg(以下、イミフィンジ)(https://www.pmda.go.jp/drugs/2018/P20180727001/670227000_23000AMX00485_A100_1.pdf (2020.2.28 確認))、スキリージ皮下注 75mg シリンジ 0.83mL (以下、スキリージ)(https://www.pmda.go.jp/drugs/2019/P20190329001/112130000_23100AMX00299_A100_1.pdf (2020.2.28 確認))、ヘムライブラ皮下注 30mg、同皮下注 60mg、同皮下注 90mg、同皮下注 105mg、同皮下 150mg(以下、ヘムライブラ)(https://www.pmda.go.jp/drugs/2018/P20180329003/450045000_23000AMX00450_A100_1.pdf (2020.2.28 確認))及びレルミナ錠 40mg (以下、レルミナ)(https://www.pmda.go.jp/drugs/2019/P20190124001/400256000_23100AMX00010_A100_1.pdf (2020.2.28 確認))の 4 品目の審査報告書のうち、審査報告(1)「非臨床薬理試験に関する資料及び機構における審査の概略」の項の全センテンスとそれに対応する AI 翻訳システムによる翻訳英文を研究対象とした。

各翻訳英文の評価は、4 段階にスコア化した(スコア 3 : このままで問題ない、スコア 2 : 完全な英文にするには一部修正が必要だが修正しなくとも意味は通じる、スコア 1 : 一部修正しないと意味が通じない、スコア 0 : 大きな問題があり大幅な修正が必要)。

また、評価の際に翻訳不備と思われた理由について、6 種類に分類した(「訳漏れ」、「原文の誤解釈」、「訳文の文法/構文」、「不適切用語」、「記号・数字」、「その他」)。

評価結果の集計及び解析は、DELL 社製 Precision 3431 SFF CTO Base において、Microsoft Word 及び Excel (Microsoft Japan) 並びに JMP 14.2 (SAS institute Japan)を用いて実施した。

2 種類の AI 翻訳システムのスコア分布の比較には、独立性の検定(累積カイ二乗検定)を用いた。

(倫理面への配慮)

本研究では動物、ヒトを研究対象としていない。

本研究においては、2 種類の AI 翻訳システムを用いているが、商品の性能比較が研究目的ではないため、商品名を匿名化し、ソフト【A】及びソフト【B】と記載する。

C. 研究結果

1) 日本語原文の 1 文当たりの文字数分布

翻訳対象となる日本語原文の特徴を抽出することを目的に、審査報告書(薬理)を構成する全センテンスの全角換算文字数を算出し、その構成割合分布を検討した。

(図 1 挿入)

その結果、イミフィンジ、スキリージ及びヘムライブラでは、25 文字超、50 文字以下のセンテンスがピークで約 35%を占め、類似した文字数分布を示していた。レルミナは、他の 3 品目とは文字数分布が明らかに異なり、25 文字以下のセンテンスが約 30%で最も多く、他の 3 品目と異なり、150 文字超の長文を含んでいた。

2) 日本語原文と翻訳後の英文センテンス数

日本語原文と翻訳後の英文センテンス数を表 1 に示す。

(表 1 挿入)

AI 翻訳システムによって翻訳された英文のセンテンス数は日本語原文のセンテンス数より多い事が示された。

イミフィンジ、スキリージ及びヘムライブラでは、ソフト【A】によるセンテンス数はソフト【B】によるものより多かったのに対し、レルミナにおいては、ソフト【B】によるセンテンス数の方が多かった。

3) 翻訳英文の評価スコア分布

翻訳英文について、4 段階にスコア化した(スコア 3 : このままで問題ない、スコア 2 : 完全な英文にするには一部修正が必要だが修正しなくとも意味は通じる、スコア 1 : 一部修正しないと意味が通じない、スコア 0 : 大きな問題があり大幅な修正が必要)。各審査品目のソフト【A】及び【B】におけるスコア分布を図 2 に示す。

(図 2 挿入)

全体的に、修正しなくとも情報として意味が通

じるスコア2以上の英文が約80%を占めていた。

ソフト【A】とソフト【B】の比較においては、ヘムライブラのみ統計学的な有意差が認められたが(独立性の検定、 $P < 0.05$)、スコア2以上の割合では差が認められなかった。傾向としては、イミフィンジ、スキリージ及びヘムライブラでは、ソフト【B】に比較してソフト【A】において、スコア3の割合が高く、スコア0の割合が低かった。レルミナに関しては、逆にソフト【B】の方がスコア3の割合が高く、スコア0の割合が低かった。

4) 翻訳英文評価における不備事項の分類

翻訳英文を評価する際に、翻訳不備と判断した事項を6種類に分類した。ソフト【A】及びソフト【B】の各審査品目における不備事項の割合を図3に示す。

(図3 挿入)

また、ソフト【A】及びソフト【B】の各審査品目における不備事項の数を表2に示す(1つのセンテンスで複数の不備事項が存在する場合もある)。

(表2 挿入)

全体的な傾向としては、ソフト【A】では「その他」が、ソフト【B】では「訳漏れ」が多く、その特徴は各審査品目によらずほぼ共通していた。

5) 翻訳不備事項の内容

スコア2、スコア1及びスコア0で翻訳不備とされた主な内容について、それぞれ表3～5にまとめた。

(表3～5 挿入)

スコア2と評価した不備事項は、情報伝達には支障がないものの一部訳漏れや用語に問題がある、上付き・下付きの表記や文法の乱れ、タイトル行の太文字が英訳に反映されていないために読みにくい等であり、修正しなくとも意味は通じるレベルの不備である。

スコア1と評価した不備事項は、このままでは日本語原文とは異なった意味に解釈されてしまうような誤訳で、部分的な修正が必須となる不備事項である。具体的には、重要な部分の訳漏れがあったり、日本語原文には存在しない数値や語句が挿入されていたり、AI翻訳による造語と思われる意

味不明の語句が含まれているような事案が認められた。また、もともとの日本語原文に省略があり、そのあいまいさから誤訳を誘発した可能性のある部分もあった。日本語特有のあいまいさから、単語間の関係性や修飾句の修飾先が間違っていたり、主語が省略されているため原文が意図したものと異なった主語に変換されるような事案である。その他に、日本にしかない固有名詞、複数の固有名詞を連結したタイトル行、元々がアルファベットの略名や単位表記で誤訳が生じていた。

スコア0と評価した不備事項は、翻訳英文が意味不明であったり、日本語原文と全く意味が異なるため、大幅な修正や、文全体の書直しが必須と判断される不備事項である。現象としては、日本語原文の途中に、注釈番号、括弧付きの引用文献、放射性同位元素ラベル表示、アルファベットの略名や単位記載がある場合に、AI翻訳システムの誤作動を誘発し、翻訳が破綻したと思われるものが多かった。また、日本語原文の構文が二重否定の場合に、片方の否定が無視され、英訳文では逆の意味になってしまうという事案や長文の日本語原文を複数の英文に分割して翻訳する過程で、誤作動を誘発したり、文章間のつながりに齟齬が生じる事案も認められた。ただし、同じようなシチュエーションで常に同じ現象が発生するとは限らず、問題なくきれいに英訳される場合と突発的に誤作動を誘発する場合があります。どのような場合に誤作動につながるのかは不明であった。その他、AI翻訳特有の問題点として、それまでの機械学習で習得した内容と思われる、日本語原文とは全く無関係の文章が突発的に挿入されたり、表の中の有効成分名が日本語原文のどこにも登場しない全く無関係の有効成分に変換されるという事案も観察された。

D. 考察

新医薬品開発のグローバル化が進む中、国際共同治験による新医薬品の世界同時開発に日本が参加することが普通となっている。それに伴い、世界で最初に日本で承認されるケースも増えてきており、日本から審査の情報をいち早く世界に発信する必要性が増してきている。特に、欧米と日本で用

法・用量が異なる場合、アジアの国にとっては、日本の審査情報が重要な役割を果たすと考えられる。今回、PMDA の新医薬品の審査報告書の英文への翻訳を効率化、高速化する手段として、最近、進歩が著しい AI 翻訳システムの実用性を評価し、利用する場合の留意点についても検討した。

最初に、日本語原文の特徴を抽出する目的で、4 品目の審査報告書(1)「非臨床薬理試験に関する資料及び機構における審査の概略」の全センテンスについて、各センテンスを構成する全角換算文字数を算出し、その分布を検討した。その結果、図 1 に示すとおり、イミフィンジ、スキリージ及びヘムライブラはほぼ類似の文字数分布を示したが、レルミナのみ異なった分布が確認された。これは、審査報告書を作成した審査担当者の個性である。

一方、2 種類の AI 翻訳システムを用いた 4 品目の審査報告書の翻訳結果を評価した。図 3 に示すとおり、ソフト【A】とソフト【B】では翻訳不備とされた事項の分類において、4 つの審査品目に抛らず、ほぼ類似のパターンを示していた。用いた AI 翻訳システムは、何れもニューラルネットワーク技術を応用した深層学習を用いたシステムであり、機械学習の過程で個性が生じるものと思われる。

次に、この審査担当者による日本語原文の個性と AI 翻訳システムの個性の相互作用について検討した。表 1 に示すように、AI 翻訳システムは、長文の日本語文を複数の英文に分割して翻訳する能力を有している。日本語原文で類似の文字数分布を示したイミフィンジ、スキリージ及びヘムライブラにおいては、ソフト【A】の方が翻訳英文数が多く、より長文の日本語原文を含んでいるレルミナにおいては、ソフト【B】の方が翻訳英文数が多くなった。これは、AI 翻訳システムにより、どのような場合に分割翻訳を行うかという判断が異なることを意味する。また、図 2 に示すように、翻訳結果スコア分布においては、イミフィンジ、スキリージ及びヘムライブラにおいてはソフト【A】の方がソフト【B】と比較してスコア 0 の割合が少なかったのに対し、レルミナでは、逆にソフト【A】におけるスコア 0 の割合が高く、長文の翻訳におい

てソフト【A】が誤作動を起こす確率が高いことが推察された。今回の検討範囲では、イミフィンジ、スキリージ、ヘムライブラにおいてはソフト【A】が、レルミナにおいてはソフト【B】が親和性を示し、相性は日本語原文の特徴に依存する可能性が考えられた。以上の結果より、今後、商用 AI システムを導入する際には、そのシステムがどのような場合に誤作動を起こす確率が高いのかを調査し、PMDA の審査報告書の特徴と照らし合わせて、どのシステムを導入すべきか検討する必要があると考えられる

続いて、AI 翻訳システムの実用性に関してであるが、図 2 に示すとおり、何れの審査品目、何れの AI 翻訳システムに抛らず約 8 割の翻訳英文については、スコア 2 以上であり、情報伝達の目的において支障はなく、AI 翻訳を利用したことを但し書きで断わっておけば許容可能と思われる英文であった。

一方、残りの約 2 割の英文については、日本語原文とは異なる情報を伝えることになるため、結局は AI 翻訳システムによって翻訳された文章全体に対して、人の手による細かいチェックと修正が必須となる。AI 翻訳システムのさらなる機械学習により、この部分を最小化することはある程度可能かもしれないが、研究結果の「5) 翻訳不備事項の内容」において示したように、日本語特有の省略等、あいまいな表現が原因と考えられる部分があり、突発的な誤作動が発生することも少なくなく、すべてを AI 翻訳に頼ることは極めて難しいと思われる。また、AI 翻訳特有の問題点として、今までに機械学習で習得したと思われる、日本語原文とは全く無関係の単語、数値、文章が突発的に挿入されたり、新たな造語や有効成分名の置換等、人の手により翻訳された英文では起こりえない事案も多く観察されており、人の手による翻訳文よりも、より注意深いチェックが必要であり、データの解釈や固有名詞等については最終的に審査担当者や審査経験者の最終チェックも必要と思われる。

ただし、スコア 1 と評価された不備事案は、多くの場合、部分的な修正で対処可能であり、翻訳文章全体の注意深いチェックにより場所が特定さえで

できれば、修正自体の手間はそれほどでもないと考えられる。問題はスコア0と評価された部分で、英訳し直しが必要となる不備事案であるが、誤作動と思われるものが多く、研究結果の「5) 翻訳不備事項の内容」において示したように、ある程度原因は推定できるため、予め対策を講じておけば最小化できる可能性がある。

審査担当者に対して、AI 翻訳システムが理解できる日本語の作成を要求することは本末転倒であるが、ある程度ルールを作ることは可能と思われる。具体的には、なるべくあいまいな表現を使用しない。日本特有の固有名詞の使用を避ける。二重否定の構文はなるべく避ける。極端に長い文章は避ける。注釈番号、引用文献は文中ではなく、文の最後に挿入する。等が考えられる。に関しては、例えば「レセプターへの結合性」ではあいまいな表現であり、「レセプターへの結合親和性」とすることにより正確な翻訳が可能になる。に関しては、例えば、試験方法の説明で、日本の検査会社が商品化した試験方法名ではなく、試験方法の簡単な原理を記載することで誤訳を避けることができる。、に関しては、状況によっては正常に翻訳される場合があるが、スコア0の誤作動を誘発する原因と推定されるため避けた方が無難である。その他に、自動翻訳を実行する前に、審査報告書の略語表の用語や、単位記載、放射性同位元素ラベル表示は翻訳しないよう辞書に登録しておくような前処理が有効と思われる。また、日本語の審査報告書では、「本薬」、「本剤」は省略語として機能するが、英文では省略語とはならず誤訳を招くため、開発記号等で予め置換する方が望ましいと思われる。

以上をまとめると、AI 翻訳システムによる翻訳英文の約 8 割はそのまま使用可能であり、人の手による翻訳文全文の詳細なチェックは必須であるものの、ある程度の対策を施すことによって、修正作業は省力化可能と思われる。方法としては、従来人の手によってゼロから翻訳していた工程を AI 翻訳システムによる翻訳文のチェック作業に置き換えれば、英文化作業のかなりの効率化、高速化が可能になると考えられる。ただし、データの解釈や

固有名詞等については、審査担当者や審査経験者の最終チェックも必要と考えられる。

最後に、以上述べたことは、薬理領域に限定した議論であり、審査報告書全体に一般化できるかどうかは明確ではない。また、今回の研究では翻訳後の英文を対象としており、翻訳過程の検証は行っていないため、どのような対策をすれば、誤訳、誤作動を予防できるかは、実際に AI 翻訳システムを動かして検討する必要があると思われる。

E . 結論

AI 翻訳システムによる翻訳英文の約 8 割はそのまま使用可能であった。残りの約 2 割の修正には、人の手による翻訳文全文の詳細なチェックが必須であるが、ある程度の対策を施すことによって、修正作業は省力化可能と思われる。方法としては、従来人の手によってゼロから翻訳していた工程を AI 翻訳システムによる翻訳文のチェック作業に置き換えれば、英文化作業のかなりの効率化、高速化が可能になると考えられる。ただし、データの解釈や固有名詞等については、審査担当者や審査経験者による最終チェックも必要と考えられる。また、実際に AI 翻訳システムを導入する際には、審査報告書の日本語文章の特徴と照らし合わせて、親和性の高いシステムを検討する必要がある。

F . 研究発表

1. 論文発表

本研究に関連した論文発表はない。

2. 学会発表

本研究に関連した学会発表はない。

G . 知的財産権の出願・登録状況

(予定を含む。)

1. 特許取得

特になし。

2. 実用新案登録

特になし。

3. その他

特になし。

表1 原文と英文翻訳後のセンテンス数の比較

	原文	翻訳英文(A)	翻訳英文(B)	表
イミフィンジ	68	72	69	1
スキリージ	29	35	31	0
ヘムライブラ	68	78	75	5
レルミナ	42	49	52	3

表2 翻訳結果評価における不備事項の数

翻訳不備の種類	イミフィンジ		スキリージ		ヘムライブラ		レルミナ	
	A	B	A	B	A	B	A	B
訳漏れ	4	30	2	17	9	47	10	25
原文の誤解釈	6	4	1	4	7	14	9	6
訳文の文法／構文	3	3	4	1	11	12	10	5
不適切用語	5	13	5	3	2	10	5	5
記号・数字	4	0	0	1	5	1	8	2
その他	12	2	6	1	26	6	15	2

表3 スコア2 の主な不備事項の内容

スコア2
原文の太文字が反映されていない。
意味は通じるが、科学論文では一般的ではない用語が使われている。
句読点が乱れている。
文章の意味は通じるが、訳されていない部分がある。
注釈番号が認識できない。
上付き、下付きが反映されていない。
単数形と複数形の乱れがある。

表4 スコア1 の主な不備事項の内容

スコア1
単語と単語の相互の関係性が間違っている(省略された日本語によるものもある)
原文には存在しない数値や単語が挿入されている。
重要な語句の訳漏れがある。
タイトル行の意味が異なる(省略された日本語によるもの)
訳文の主語が間違っている(省略された日本語によるもの)
英語としても一般的ではない意味不明な造語がある。
修飾句が修飾する先が間違っている。
日本にしかない固有名詞のため、英語に直訳すると異なる意味になる。
原文のアルファベット略名や単位記載を更に変換し、意味不明になる。

表5 スコア0 の主な不備事項の内容

スコア0
文章の意味が異なるため、大幅に修正が必要
二重否定が訳されていないため、逆の意味になっている。
原文のどこにも登場しない有効成分名に置換されている。
原文のかなりの部分が訳されていない。
原文の中にアルファベットの略名や単位記載があると、誤作動を誘発し英訳できなくなる。
長い原文を分割して英訳する際に、文章間のつながりに齟齬を生じる。
長い原文を分割して英訳する過程で誤作動を誘発し、意味不明な訳文となっている。
数値の挿入箇所が間違っているため、原文と英訳文で試験結果が異なる。
原文の文章途中に()で引用文献が挿入されていると、誤作動を誘発し、英訳できなくなる。
原文とは全く無関係の文章が挿入されている。

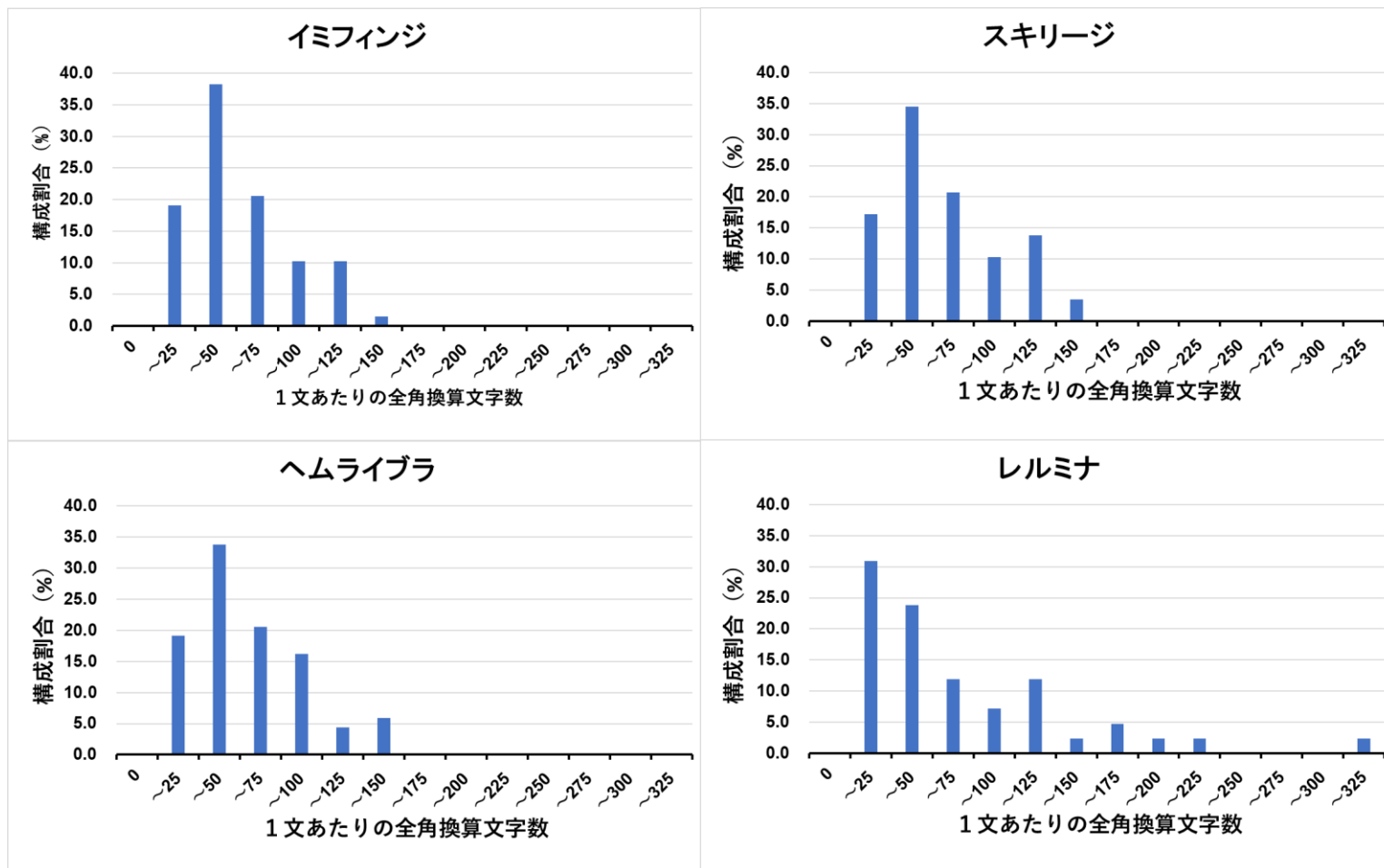


図1 原文の1文あたりの全角換算文字数の分布

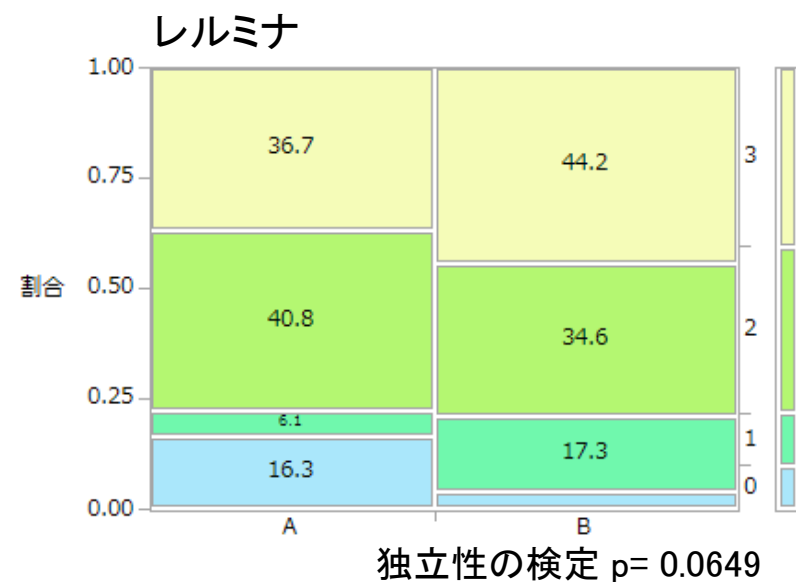
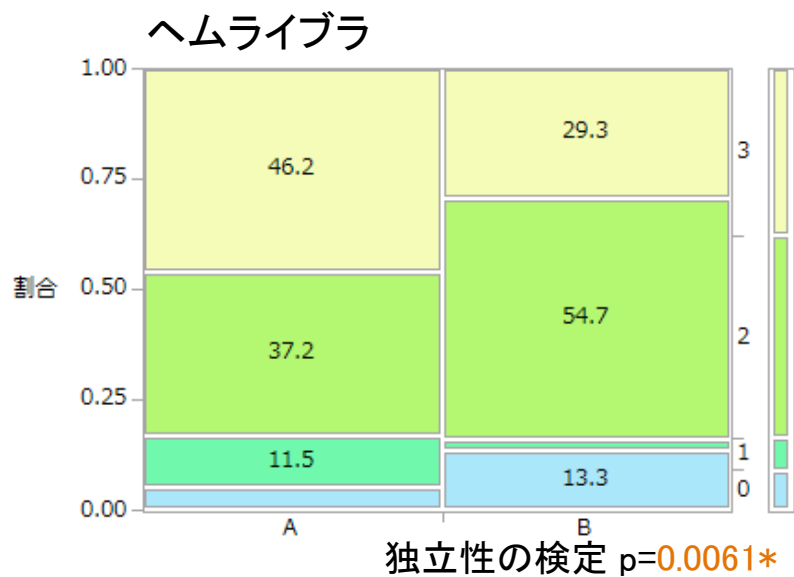
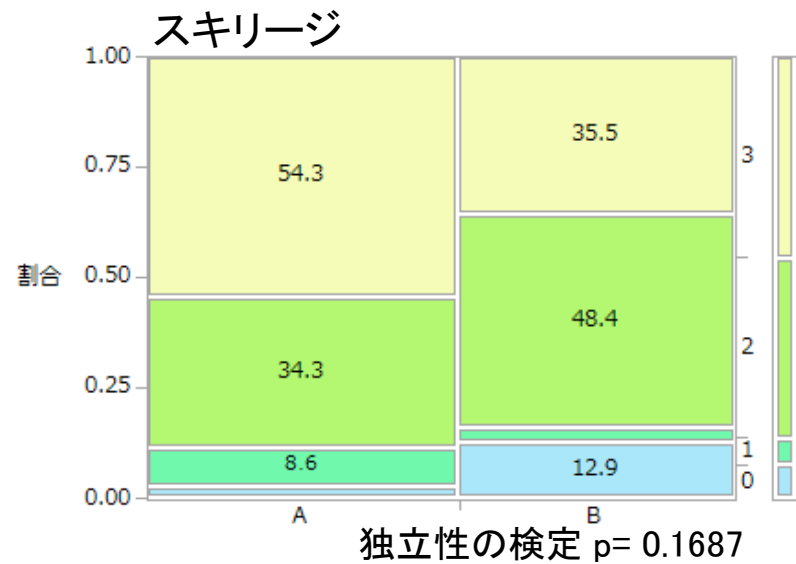
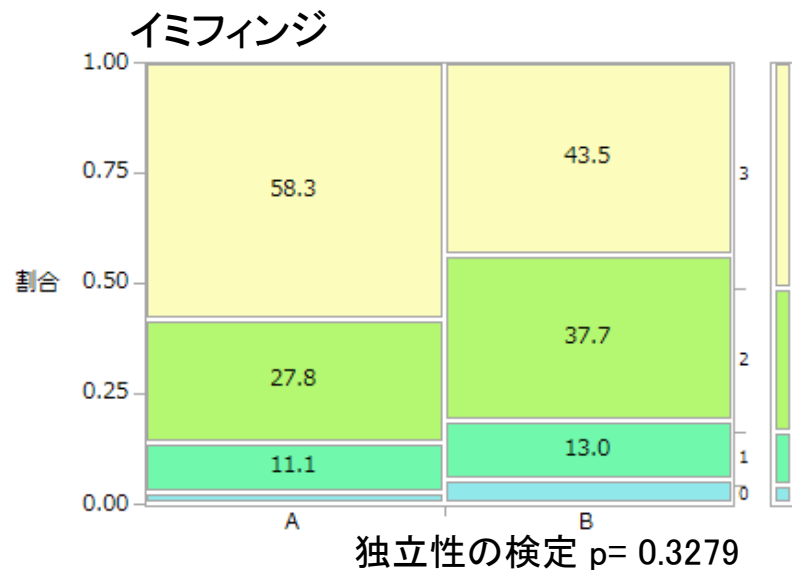


図2 翻訳ソフトAとBによる翻訳結果スコア分布の比較

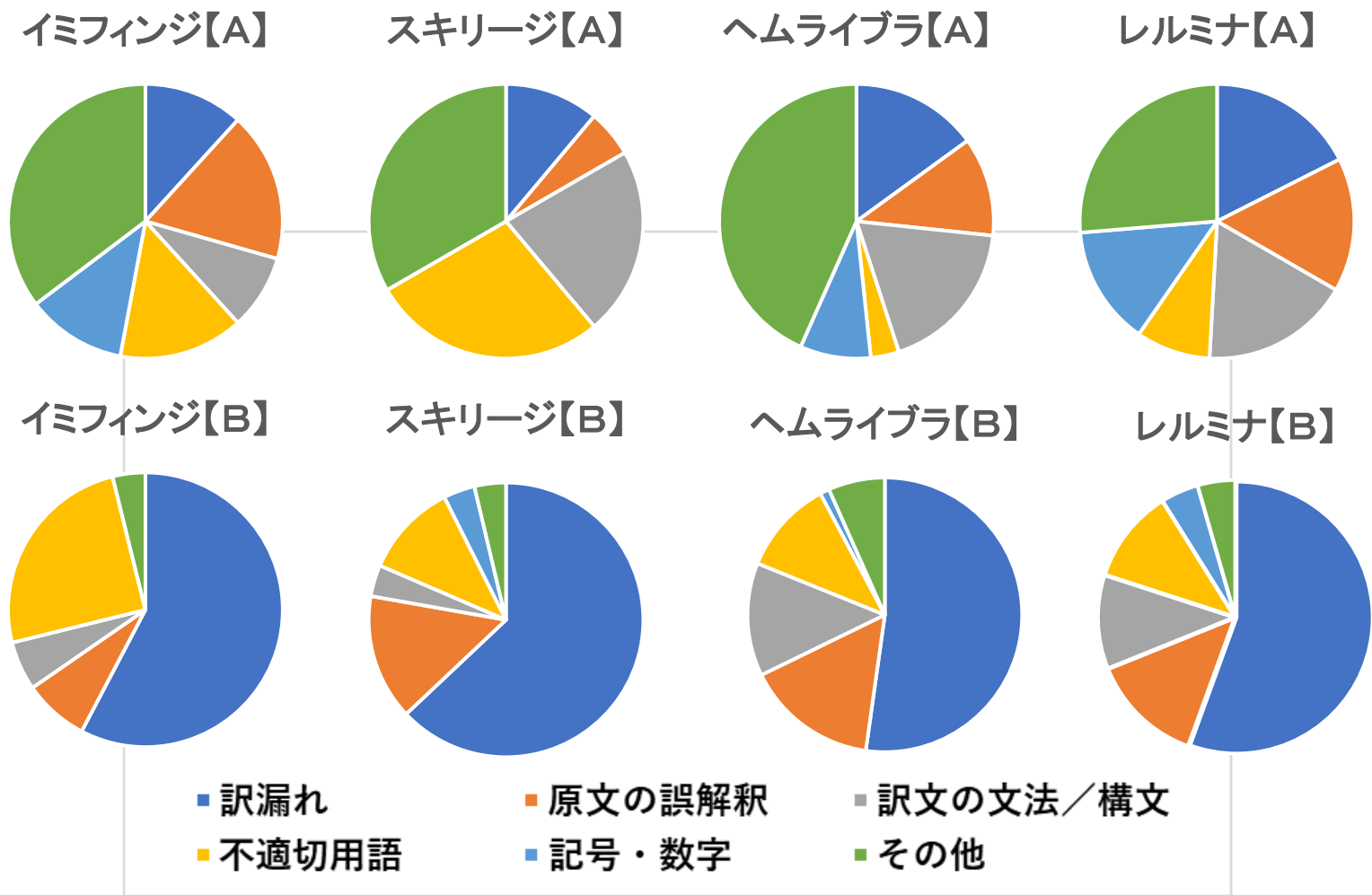


図3 翻訳結果評価における不備事項分類の割合比較

別添 5

研究成果の刊行に関する一覧表

書籍

著者氏名	論文タイトル名	書籍全体の編集者名	書 籍 名	出版社名	出版地	出版年	ページ
	該当なし						

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
	該当なし				

< 注意 >

分担研究報告書の F . 研究発表に記載したものは、研究成果の刊行に関する一覧表にもご記載ください。
また、その刊行物又はその別刷一部を添えてご提出ください。

別添 6

厚生労働科学研究費における倫理審査及び利益相反の管理の状況に関する報告

(作成上の留意事項)

厚生労働科学研究費における倫理審査及び利益相反の管理の状況に関する報告は、「厚生労働科学研究費における倫理審査及び利益相反の管理の状況に関する報告について(平成26年4月14日科発0414第5号)」の別紙に定める様式の写しを添付するが、別紙 5 に当該様式を再掲するので参考にすること。

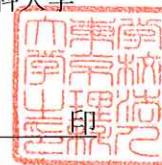
2020年5月20日

厚生労働大臣 殿

機関名 学校法人東京理科大学

所属研究機関長 職名 理事長

氏名 本山 和夫



次の職員の令和元年度厚生労働行政推進調査事業費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

- 1. 研究事業名 厚生労働科学特別研究事業
- 2. 研究課題名 医薬品評価情報作成システムの有用性評価研究 (19CA2017)
- 3. 研究者名 薬学部薬学科・教授
鹿野 真弓 (シカノ マユミ)

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
ヒトゲノム・遺伝子解析研究に関する倫理指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
人を対象とする医学系研究に関する倫理指針 (※3)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること (指針の名称:)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他 (特記事項)

(※2) 未審査の場合は、その理由を記載すること。

(※3) 廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> (有の場合はその内容:)

(留意事項) ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

2020年 2月 19日

厚生労働大臣 殿

機関名 武蔵野大学

所属研究機関長 職名 学長

氏名 西本 照真



次の職員の令和元年度厚生労働行政推進調査事業費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

1. 研究事業名 厚生労働科学特別研究事業
2. 研究課題名 医薬品評価情報作成システムの有用性評価研究 (19CA2017)
3. 研究者名 (所属部局・職名) 薬学部・教授
(氏名・フリガナ) 永井 尚美・ナガイ ナオミ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
ヒトゲノム・遺伝子解析研究に関する倫理指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
人を対象とする医学系研究に関する倫理指針 (※3)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること (指針の名称:)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他 (特記事項)

(※2) 未審査に場合は、その理由を記載すること。

(※3) 廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> (有の場合はその内容:)

(留意事項) ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和2年2月25日

厚生労働大臣 殿

機関名 横浜薬科大学

所属研究機関長 職名 学長

氏名 江崎 玲於奈



印

次の職員の令和元年度厚生労働行政推進調査事業費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

1. 研究事業名 厚生労働科学特別研究事業

2. 研究課題名 医薬品評価情報作成システムの有用性評価研究 (19CA2017)

3. 研究者名 (所属部局・職名) 薬学部 臨床薬学科 レギュラトリーサイエンス部門 教授

(氏名・フリガナ) 山田 博章 (ヤマダ ヒロアキ)

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
ヒトゲノム・遺伝子解析研究に関する倫理指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
人を対象とする医学系研究に関する倫理指針 (※3)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること (指針の名称:)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他 (特記事項)

(※2) 未審査に場合は、その理由を記載すること。

(※3) 廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> (有の場合はその内容:)

(留意事項) ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。