

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