

日本医療研究開発機構 医薬品等規制調和・評価研究事業 事後評価報告書

公開

I 基本情報

研究開発課題名: (日本語) 医薬品安全性評価の迅速化・高精度化に資する非臨床試験毒性用語の標準化とデータベース構築

(英語) Construction of non-clinical toxicity study terms standardized toward accelerate and high-accuracy of drug safety evaluation

研究開発実施期間: 令和 4 年 4 月 1 日～令和 7 年 3 月 31 日

研究開発代表者 氏名: (日本語) 高橋 祐次

(英語) Yuhji Taquahashi

研究開発代表者 所属機関・部署・役職:

(日本語) 国立医薬品食品衛生研究所 安全性生物試験研究センター 毒性部 動物管理室 室長

(英語) Section Chief, Animal Care Section, Division of Cellular and Molecular Toxicology, Center for Biological Safety & Research, National Institute of Health Sciences

II 研究開発の概要

研究開発の成果およびその意義等

和文:

本研究は、1994 年に国立医薬品食品衛生試験所安全性生物試験研究センター（当時）が中心となって編集・発刊された「毒性試験用語集」を基に、毒性試験で用いられる用語の標準化を行い Web 検索が可能なデータベース (DB) を構築することを目的とした。日本製薬工業協会 (JPMA) の調査によれば、「毒性試験用語集」は非臨床試験領域で最も参照されてきた辞書である。毒性試験用語は毒性研究の基盤資産であり、正確かつ一貫性のある記録と利活用が求められる。しかしながら、同書は発刊から 30 年以上が経過しているにもかかわらず一度も改訂が行われていない。その結果、毒性試験担当者並びに申請資料作成者からは現行の実務で使用される用語との乖離が懸念されていた。加えて、同書は紙媒体であるため、情報の更新・共有・管理といった点にも課題があった。

医薬品の開発プロセスにおいて、臨床試験を安全に実施するためには非臨床安全性試験は欠かせない評価ステップであり、その過程では極めて大量のデータが生成される。たとえば、標準的なラット 13 週間反復投与試験においては、1 匹あたりおよそ 550 項目にわたるデータが記録され、1 試験全体ではおよそ 110,000 件以上に及ぶデータが生成される (Horii I, EHP Supp, 1994)。大量データの取り扱いの

信頼性を確保し、試験の再現性および再構築性を担保するため、1980 年代には GLP (Good Laboratory Practice) 基準が制定された。そして、これら多くのデータの収集、一元管理および統計解析を実施するためコンピュータを利用したデータ管理支援アプリケーションソフトウェアが利用されている。毒性試験では一般状態、体重、摂餌量、血液検査などの定量的・定性的データから構成され、これらは個体ごとに整理され、統計解析後に帳票として最終報告書に組み込まれる。しかしながら、試験施設ごとに使用しているアプリケーションソフトウェアが異なるため、データ構造の互換性がなく、試験横断的に活用することが困難な状況となっている。医薬品承認申請資料は、国際会議 ICH で定められた Common Technical Document (CTD) に準拠して作成される。近年は eCTD (electronic CTD) による電子提出が一般化しており、資料作成、審査、保管といった作業の効率化が進んでいるもののデータ解析に適した構造にはなっていない。

こうした背景のもと、米国食品医薬品局 (FDA) は、治験開始前に必要な IND (Investigational New Drug) 申請に際し、非臨床試験データを効率的に解析できるようにするため、2016 年より SEND (Standard for Exchange of Nonclinical Data) の導入を義務化した。SEND は、CDISC (Clinical Data Interchange Standards Consortium) が策定した標準データ形式であり、非臨床データの電子的提出と二次利用を可能にする国際的な枠組みである。しかしながら、SEND においても一般状態所見などの自由記述的なデータ領域については、未だ標準化された Control Terminology (CT) が整備されておらず、各試験施設が独自の用語を使用しているのが実情である。このことは、統合的な解析やリスク予測に大きな障害となっており記録された一般状態観察が示す病態の本質が見落とされる可能性がある。

実際、2016 年に実施された BIA 10-2474 の FIH 試験では、被験者に不可逆的な中枢神経障害や死亡が発生するという重大な事態が起こった (Butler D & Callaway E, Nature 2016)。事後の解析では、原因となった分子機序の特定には至っていないものの、非臨床段階での所見記録において用語の不統一がリスクの見落としの原因である可能性が指摘されている (Hayes AW et al., Crit Rev Toxicol. 2021; 西村他, 日本獣医史学雑誌 2024)。臨床領域では副作用・有害事象の記録に関しては MedDRA (Medical Dictionary for Regulatory Activities) という国際的に標準化された辞書が活用されており、これにより国・機関間での情報の共有や解析が可能となっている。非臨床領域においても同様の用語整備が求められているが、これまでに整備された体系的な辞書は存在しないため共通したデータ解析プラットフォームでの再利用ができない状況であり、昨今、飛躍的に進歩によりあらゆる分野での利用が進んでいる AI 解析の足かせとなっている。

こうした課題を克服するため、本研究ではまず、1994 年版「毒性試験用語集」の全項目 (約 500 頁) のデジタル化を行い、データの共有・編集・改訂作業が可能なデータとした。JPMA および安全性評価研究会の協力により、現行の試験報告書で用いられている用語の実態把握のため、PMDA の医療用医薬品情報検索データベースを用いて 1999 年～2020 年に承認された 1945 品目の CTD m2.6.6 に記載された一般状態症状用語 (日本語) の分析・抽出を行った。用語の解析にはテキストマイニングの手法も併用した。これらの解析結果を踏まえて、約 2,300 語を新たに追加した。

データベース化には、情報を体系的に整理し、検索・利活用可能な構造で保存することが必要となる。そのために必要となるのが、オントロジー (Ontology) である。オントロジーは用語の意味、階層、関係性を定義する枠組みである。これにより、用語の一貫性や解釈の明確化が可能となり、リードアクロスや AI 解析などへの利用が促進されることが、オントロジーはデータベースの機能を高度化する鍵となる。具体的には、Main Category Term (1994 年版毒性試験用語集で用いられている A~T の分類)、Primary Group Term (高次のグループ分類)、Secondary Group Term (低次のグループ分類)、Basic Term (基本的に使用される用語として定義) 及び Equivalent Term (Basic Term と同様の意味で使用

される用語)として整理しデータベース化した。毒性病理用語に関してはミクロ所見について診断基準国際統一化(International Harmonization of Nomenclature and Diagnostic Criteria (INHAND))が進められているが、肉眼所見については標準化されていないことから、日本毒性病理学会(JSTP)の国際用語委員会の協力により、毒性試験用語集の剖検所見を基に肉眼所見に関し、一般所見、臓器特異的所見、外観、その他について SEND 用語の区別を示したリスト(和・英)を作成し、毒性病理学会の HP に公開した。また、げっ歯類の一般毒性試験用及びがん原性試験用常用所見並びに、イヌの一般毒性試験用常用所見についても SEND 及び INHAND との国際統一性を考慮したリスト(英)を作成し公開した。

申請資料や報告書等に使用される「規制関連用語」について PMDA との連携体制の下で用語の抽出作業を進め、日本語と英語の対訳を含めた形で整理を行った。

これらのデータを一般公開しパブリックコメントを求めるため、ユーザー登録、検索機能、ユーザーインターフェースデザインを含むウェブベースの利用環境を整備し、「新毒性試験用語集」としてプロトタイプを公開するに至った。ウェブベースのデータベースの構築には、不正アクセスや情報漏洩を防止するための堅牢なセキュリティ対策が不可欠であることから、これらの専門技術を有する東京都ビジネスサービスに委託して設計を行った。

「新毒性用語集」のウェブサイトに関して、各方面にアナウンスを行いパブリックコメントへの協力を求めた。約3ヶ月間のパブリックコメントの期間において10件のコメントが寄せられた。いずれも、利便性の高い毒性試験用語集のデータベースとして発展させるため重要な示唆を含んでいる。

現状では、非臨床試験領域において国際的な統一辞書は存在せず、日本が主導する形でグローバルな用語整備を推進することが期待される。ICH 参画を通じた国際的なネットワークも本研究の成果活用に資することができる。本データベースが直ちに国際標準化に至るものではないが、各種非臨床試験における用語整備や SEND 入力ルール策定の支援する基盤となることが期待される。特に、SEND において CT が未設定の一般状態所見領域において、入力標準化と用語の明確化に貢献することが想定される。医薬品開発における非臨床安全性試験の最終的な目標は、臨床試験に参加する被験者の安全性を確保することである。毒性評価に用いられる一般状態用語が統一され、データベースとして活用可能になることで、AI や機械学習を用いた毒性影響のパターン解析が可能となり、クラスエフェクトやドラッグエフェクトの抽出、予測的評価に活用できる。

なお、本研究で整備した一般状態観察用語は日本語であり、英語に関しては見直しが進んでいない。昨今は ChatGPT 等の生成系 AI を用いた翻訳支援も可能となっているが、本分野においては日英対訳の教師データ(AI に学習させる正しいデータセット)が十分に整備されておらず、毒性試験用語の分野特有の訳語選定には一定の限界がある。実際に、ChatGPT4o での翻訳を試みたが十分な結果が得られなかった。

1994 年の毒性試験用語集は、コンピュータでの活用が難しいながらも、我が国の非臨床評価の基盤資産として長年用いられてきた。本研究では、その電子化とデータベース化を実現し、今後の研究・実務の橋渡しとなる成果を得た。一方で、用語は科学の進展とともに変化するものであり、継続的な見直しと運用体制の構築が求められる。これらの活動を通じて、日本の創薬研究基盤の強化と国際的な医薬品安全性の向上に資することを目指すものである。以上の取り組みにより、非臨床毒性評価における所見記録の一貫性・比較性・国際的な整合性が向上し、リードアクロスやデータ駆動型評価手法(NAMs)の信頼性向上にも資する。動物実験の価値を維持・拡張しつつ、将来的な毒性評価の高度化と効率化を支える基盤整備となることが期待される。

英文：

This project was initiated to standardize terminology used in toxicity studies and to develop a searchable, web-based database, based on the, “*Dokuseishiken yougoshyou*”, Toxicological Terminology Book published in 1994 by Japan’s National Institute of Health Sciences. Despite being the most referenced resource in Japan’s nonclinical studies, the book had not been revised in over 30 years, leading to inconsistencies between actual practice and documented terminology. Its paper format also hindered timely updates and digital integration. Accurate and consistent use of terminology is essential in nonclinical safety studies, which generate large volumes of data. A standard 13-week rat toxicity study can produce over 110,000 data points. To manage such data, Good Laboratory Practice (GLP) standards and digital tools have been widely adopted. However, variation in software systems across testing facilities limits interoperability and cross-study analyses. While regulatory submissions now follow the ICH’s Common Technical Document (CTD) and its electronic counterpart (eCTD), these formats are not optimized for data reanalysis. To address this, the U.S. FDA mandated the use of SEND (Standard for Exchange of Nonclinical Data) in 2016. Although SEND provides standardized formats, it lacks defined Control Terminology (CT) for free-text fields like general condition observations. This omission results in inconsistent language use and hampers integrated risk analysis. A notable example of the consequences of inconsistent terminology is the 2016 BIA 10-2474 first-in-human trial, which resulted in severe CNS damage and fatality. Post-trial analyses suggested that inconsistent preclinical terminology may have contributed to the failure to identify risk signals. In contrast, the clinical field benefits from MedDRA, a globally standardized dictionary that enables cross-border sharing and analysis of adverse event data. A comparable resource is lacking in the nonclinical field, limiting the use of AI and machine learning for predictive toxicology.

To address these issues, this study digitized all entries (approx. 500 pages) from the 1994 book. With support from JPMA and the Safety Evaluation Forum, current terminology used in test reports was analyzed using PMDA’s drug information database. General condition terms from CTD m2.6.6 of 1,945 approved drugs (1999–2020) were extracted and text-mined, adding about 2,300 new terms. Database development required structuring the information for search and reuse, for which ontology was introduced. Ontology defines the meaning, hierarchy, and relationships of terms, enabling consistency and interpretability, and facilitating AI-based analyses. Terms were categorized as Main Category Term, Primary Group Term, Secondary Group Term, Basic Term, and Equivalent Term. For pathology findings, macroscopic terms were reviewed with the Japanese Society of Toxicologic Pathology. Regulatory terms used in applications and reports were extracted in collaboration with PMDA and translated into English. To gather public feedback, a prototype of the “New Toxicological Terminology Database” was launched online (<https://test.cbsr-yougo.org>), featuring user registration, search functions and a user-friendly interface. Security measures were implemented in collaboration with Tokyo Business Service to prevent unauthorized access and data leakage.

Public comments were made, resulting in 10 suggestions during a 3-month period, all of which provided valuable insights into improving the database’s usability. Currently, no internationally unified dictionary exists for nonclinical studies. Japan is expected to take a leading role in global

terminology standardization, leveraging ICH networks.

While the database is not yet an international standard, it is expected to serve as a foundation for terminology alignment and SEND rule development. In the long term, this initiative aims to enhance the consistency, reusability, and international compatibility of preclinical toxicity data—supporting next-generation, data-driven safety assessments and contributing to the reduction and refinement of animal testing.