

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) 医療機器不具合用語集のシグナルディテクションへの適用とそれを支援するツール群の研究開発

(英語) Study of Medical Device Adverse Event Terminology Application to Signal Detection and Supporting Tools development

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) 横井 英人
(英語) Yokoi Hideto

研究開発代表者 所属機関・部署・役職:
(日本語) 国立大学法人 香川大学・医学部附属病院 医療情報部・教授
(英語) Kagawa University Hospital・Medical Information Department・Professor

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:

<背景>

医療機器の安全対策の情報化は、医薬品に比べやや遅れている感があったが、本邦に於いて 2020 年に不具合用語集の用語が不具合報告システム (XML 作成ツール) 上で選択できるようにシステム化されたことで、データのコード化、標準化が進んだ。

これらの情報化がなされると、市販後に得られる安全情報 (副作用報告・不具合報告など) により、対象となる品目の安全プロファイルが治験で得られたときよりも明確化することで、より有効な安全対策をとることができる。

シグナルディテクション (SD) は薬事に於ける安全対策の手法の一つとして、医薬品では広く行われており、本邦でも医薬品医療機器総合機構 (PMDA) が 2009 年から導入している。医療機器に於いては、欧米では実施例の報告があるが、本邦ではまだ行われていない。

前述したように、日本に於いて医療機器の不具合用語集が活性化し、実際にデータベースの中で統一的な用語により検索が可能となったのは 2020 年からであることから、大規模なデータベース検索の検討はこれからであると言える。

国際医療機器規制当局フォーラム IMDRF (International Medical Device Regulators Forum) では、様々

な国際整合作業が行われており、本研究ではこれらの国際的な安全対策の規制手法と整合が取れる手法を案出する必要がある。一方、日本の規制制度下で日本語を使用した報告様式から、迅速に医療機器の安全対策に資する情報を得るための日本独自の手法も必要となり得ることから、国際整合と国内運用を両立するという課題を解決することが望まれていると考える。特に近年、人工知能を背景とした自然言語処理技術の発展が加速しており、それらの技術を踏まえた新たな解析手法の導入が期待される。

<研究の目的>

本研究の目的は2つある。一つは、これまで構築された医療機器不具合用語集および医療機器不具合報告システムの実運用状況を調査し、そこから得られたデータを用いた医療機器に於けるシグナルディテクション（以下SD）を行うことである。具体的には医薬品医療機器総合機構（PMDA）で既に行われている医薬品に対するSDと同様の手法の医療機器への適用の可否及び医療機器に合わせた変更の必要性に係る検討、また海外で行われた医療機器に対するSDの実情調査と、それを元に立案したSDプロトコルの実行である。この研究期間内では、SDプロトコルを実行した結果について、検出されたシグナルをレトロスペクティブに評価し、その妥当性を検討すると共に閾値の最適化に関する知見を得る。そして、それに基づいたSD検出用モデルを準備し、実装用の仕様案をまとめることとする。これにより、PMDAでのSD実施の方向性が定まり、国内に於ける医療機器安全対策の効率化、省力化を図ることがねらいとなる。

もう一つは、2020年から本格運用フェーズに入った医療機器不具合用語集の維持管理に関する技術的・学術的支援を行い、また判明した運用上の問題点を解決する新たな維持管理手法を案出することである。本邦に於ける医療機器不具合用語集は、一般社団法人日本医療機器産業連合会（医機連）にて医療機器グループ毎に専用の用語集を日本語で作成した日本独自の用語集（以下、医機連用語集）と、近年IMDRF（International Medical Device Regulators Forum）で作成された不具合用語集（以下、IMDRF用語集）の翻訳版を合成して編集されたものである。この運用と国際整合性の過程で、IMDRF用語集の翻訳とのマッピング（対応表作成）作業が完了しているが、用語集の改訂への対応のため、その作業は今後も継続する。今後、国際的にSDが行われたときの対応を含めて、状況の調査と共に、これからの用語集のあり方と維持管理手法の検討を行うこととする。また、国内で既に蓄積されている、過去の不具合報告のうち、用手的に不具合事象名を設定していた時期のデータについて、最新の不具合用語集へのマッピングするデータマネジメントを行い、それによりシグナル検出精度の向上が見られるか、検討する。SDの精度は、蓄積される情報の正確性（正確に報告されているか）と、適切性（適切な粒度で報告されているか）の2点に影響を受ける。特に適切性は用語集がSDに適切な粒度で作成されているかに依存するため、本研究で得たSDに必要な用語集粒度に関する情報のフィードバックを行う。特に重要な分野を絞り込み、用語定義の精緻化を行い、SD精度の変容について評価しながら、よりよい用語集編集の検討を行う。これらの結果、本邦の不具合報告DBを用いたSDが国際的に規制を行う際にも有効に利用することをねらいとする。

<結果・考察>

SDについては、医薬品に於けるSDの現状を確認し、医療機器への応用の検討を行った。医薬品は副作用という一つの属性のみで安全情報が完結するのに対して、医療機器では（機器の）不具合と（患者の）健康被害という二つの属性があり、その扱いの難しさから不均衡分析に関する評価は難航した。

研究班では、不均衡分析以外の手法として「報告数の増加＝シグナル」という定義のもとに、変化点分析（Change Point Analysis）手法の妥当性を調べた。PMDAから「適正使用のお願い」が出された時期との関連性を検討がなされた。大動脈用ステントグラフトのエンドリークについて検討したところ、信頼性の推移が90%を超えた時点を指標としたときに「お願い」が出される1年前に増加を検出、また変化点の時点を指標としたときには2年前に増加を検出できた。最終年度には計画通りに改訂されたプロトコルをもとに前年とは違う評価対象である埋め込み型心臓ペースメーカーの健康被害として再手術を検討した。その結果として、再

手術の報告数の推移は時系列的に増加傾向にあって、その変化点としては2020年8月となり、信頼性としては100%であった。不均衡分析の1手法であるPRR法については、2年度の大動脈用ステントグラフトに加え、一般的名称「植込み型心臓ペースメーカ」に対して実施した。回収・改修事例を含む製品×事象のPRR値に着目すると、時系列ごとの解析においては回収・改修の実施時期に対応した値の増減が認められた。最終年度にはRのライブラリPhVIDを用いて、ROR,PRR以外のGPS,BCPNNといった不均衡分析手法についても実施した。医療機器は、医薬品と異なり、それぞれの機器で発生する不具合、健康被害が明らかに違う（電源の有無により、電源の不具合がある医療機器とない医療機器に分かれる）という特性がある。したがって同じような安全プロファイルを持った医療機器のグループ内で分析を行う必要がある。このような背景から、副作用報告の分析手法である不均衡分析をそのまま医療機器の不具合報告で適用するのは適切ではないとされてきた。しかし、実際に不均衡分析を試みた報告はないため、本研究で検証を試みた。分析対象をペースメーカの【不具合】事象名の報告とし、2021年4月から6月までに報告された440件とした。シグナル有りと判定されたのはPRR：94件、POR：89件、GPS：40件、BCPNN：21件であった。4つの分析すべてでシグナル有りと判定されたのは17件あった。GPSでは6件が単独でシグナル有りと判定されていたが、BCPNNでは単独で判定されたものではなく、必ず別の判定結果でシグナル有りとなっているものが含まれていた。この結果より、PRR、PORでは過剰な検出がされる傾向があるが、ベイズ理論を用いた補正が行われるGPSやBCPNNではより絞り込まれた検出結果となった。

用語集の維持管理については、技術的支援として、医機連不具合用語WGとともに不具合用語集第6版改訂時に発生した課題の検討を行った。用語集の維持管理作業は、他の用語集の維持管理などと比較して、一般的な作業内容を行っていることが判明したが、実際に発生した課題は、本用語集に特異的な内容が多く、システム的な対応には今しばらく検討が必要とみられた。学術的支援としてはIMDRFが公開している医療機器不具合用語集(IMDRF用語集)の機械翻訳について種々のモデルを検討し、GPT-4が最良であることが明らかとなった。また自動マッピングツールの性能評価を継続した。またオントロジー（意味論）理論に基づき用語集間の不整合検出を行うツールについてPythonにて一連の作業を実施できるよう構築した。用語集維持管理の新手法検討については、技術的支援・学術的支援の運用として、香川大学が整備・管理するハンドリングシステムへの機能追加を実施し、作業者の技量・時間に頼らずに、基礎能力があれば誰でも維持管理可能なシステムへとバージョンアップすることができた。

<総括>

SDについては、医薬品と医療機器の相違点を把握するのに一定の時間を要したが、医療機器の特殊性を明らかにすることができた。それを受けて医薬品で頻用される不均衡分析以外の手法の妥当性を検討した。医療機器の特性に合わせた明快な手法を案出するまでは至らなかったが、有用性についていくつかの示唆を得ることができた。

用語集の維持管理については、ツールの安定的運用が概ね実証され、今後の運用の継続性の見込みは立ちつつあるという感触を得た。

英文：

<Background>

The informatization of safety activity of medical devices (MD) has lagged behind that for pharmaceutical products. In Japan, however, the adverse event (AE) terminology was systematized in 2020 so that the terms in the AE reporting system (XML creation tool) can be selected, and the coding and standardization of data has progressed.

Once these informatization efforts are made, safety information obtained post-marketing (e.g., adverse drug reaction reports, AE reports for MD, etc.) will clarify the safety profile of the subject item more clearly than when obtained in clinical trials, allowing more effective safety actions to be taken.

Signal detection (SD) is widely used in pharmaceuticals as one of the safety measures in pharmaceutical affairs, and was introduced in Japanese regulatory action by the Pharmaceuticals and Medical Devices Agency (PMDA) in 2009. In the case of medical devices, there have been reports of cases of implementation in Europe and the United States, but not yet in Japan.

<Purpose of the Study>

The purpose of this study is twofold. The first is to investigate the actual operation of the medical device AE terminology and medical device AE reporting system that have been established so far, and to perform signal detection (SD) for medical devices using the data obtained from these systems. The second is to provide technical and academic support for the maintenance of the MD AE terminology, which has entered the full operational phase in 2020, and to propose new maintenance methods to solve operational problems that have been identified.

<Results and Discussion>

The research team confirmed the current status of SD for pharmaceuticals and examined its application to MD. While safety information for pharmaceuticals is completed with only one attribute, adverse drug reactions, MD have two attributes, namely, device problem and patient problem, and the difficulty of handling these attributes made the evaluation of disproportionality analysis (DA) difficult. The team examined the validity of the Change Point Analysis method based on the definition that “an increase in the number of reports = a signal” as a method other than DA, and examined the relevance to the period when the PMDA issued a “request for proper use”. When the reliability trend was over 90%, an increase was detected one year before the “request” was issued.

The PRR method, a method of DA, was conducted for the generic name “implantable cardiac pacemaker” in addition to aortic stent grafts in 2nd FY. Focusing on the PRR values for [product] x [event] including cases of recall and repair, an increase or decrease in values corresponding to the timing of recall and repair was observed in the analysis for each time series. In the final year of the project, we introduced the R library PhVID to conduct DA methods such as GPS and BCPNN in addition to ROR and PRR. Medical devices, unlike pharmaceuticals, have the characteristic that device problem and patient problem that occur with each device are clearly different (depending on the presence or absence of a power supply, there are two types of medical devices: those with power supply failures and those without power supply failures). Therefore, analysis should be conducted within a group of medical devices with similar safety profiles.

For the maintenance of the MD AE Terminology, as technical support, the issues that arose during the revision of the 6th edition of the AE terminology were examined together with the WG of the MD AE

terminology. It was found that the maintenance and management of the terminology was more general than the maintenance and management of other terminologies, but many of the issues that actually arose were specific to this terminology, and it is necessary to consider systemic measures for a while now. For academic support, various models for machine translation of the IMDRF MD AE terminology (IMDRF terminology) were examined, and GPT-4 was found to be the best. We also continued to evaluate the performance of automatic mapping tools. We developed a tool for detecting inconsistencies between terminologies based on ontology (semantics) theory, which can perform a series of tasks in Python. As for the study of a new method of glossary maintenance, we added functions to the handling system maintained by Kagawa University as an application of technical and academic support, and were able to upgrade the system to a system that can be maintained by anyone with basic skills, without relying on the skills and time of the workers.