

令和4年度 医療情報データベース推進コンソーシアム
報告書

令和5年3月31日

はじめに

医薬品の製造販売後の調査に医療情報データベース（診療録等の情報を電子計算機を用いて検索することができるように体系的に構成したもの）を利用した際の再審査及び再評価の申請書に添付する資料の信頼性を確保するため、平成 29 年 10 月 26 日に「医薬品の製造販売後の調査及び試験の実施の基準に関する省令等の一部を改正する省令」（平成 29 年厚生労働省令第 116 号。以下「改正 GPSP 省令」という。）が公布され、平成 30 年 4 月 1 日から施行されている。改正 GPSP 省令では、これまでの使用成績調査が医療機関から直接収集する情報を用いる調査と規定されるとともに、新たに製造販売後データベース調査（以下「製販後 DB 調査」という。）の定義が設けられ、製販後 DB 調査が医療情報 DB を用いた調査であることが明示的に規定された。製販後 DB 調査に関連するガイドライン等が発出され、当該調査を実施するための環境は整備されつつあるものの、製販後 DB 調査を実施した件数は伸び悩んでおり改善の余地は大きい。そこで、令和 3 年度厚生労働省予算の主要事項の一つとして、医療情報 DB の活用推進に向けた環境整備を挙げ、製薬企業や医療情報 DB 事業者等と利活用情報やアウトカム定義※等について検討・共有することを目的としたコンソーシアムを設置し、アウトカムバリデーションについての現状課題や先行事例の共有をおこない、今後の検討事項について取りまとめ、公表した。

昨今、国内におけるアウトカムバリデーション研究の報告も増えてきていることから、令和 4 年度は、アウトカムバリデーション研究のさらなる発展を目指し、国内等におけるアウトカムバリデーション研究事例の共有、アウトカムバリデーションの実施方法や効率的な実施に向けての課題等を取りまとめることとした。

※データベース調査を実施する際に、目的とする有害事象（＝アウトカム）を特定するために必要とされる抽出条件

アウトカムバリデーション手法の整理

(1) 国内外のアウトカム定義とバリデーションの比較

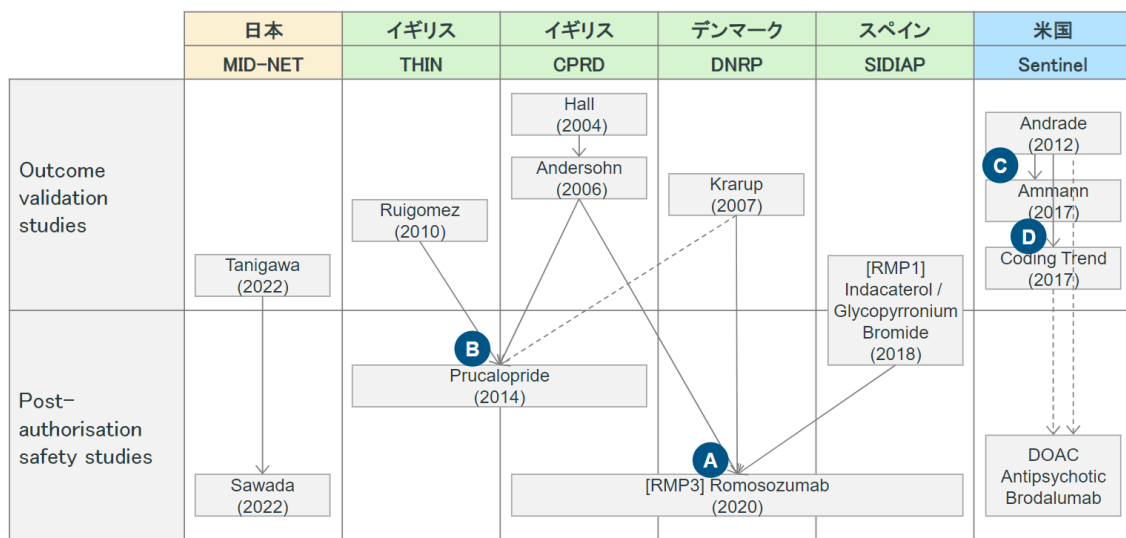
アウトカムの定義とバリデーション手法の国内外での差異を確認するため、日本・欧州・米国の事例を検索して比較した。

対象アウトカムは、日本でも既にバリデーション研究結果が公表されている脳梗塞・感染症・悪性腫瘍を選び、海外の実例は令和3年度コンソーシアムの調査にて活用したWEBサイト「EU PAS Register」と「Sentinel Initiative」で検索した。

1. アウトカムバリデーション研究結果の再利用状況（例：脳梗塞）

脳梗塞をアウトカムとしたバリデーション研究と PASS（Post-authorisation Safety Studies）の関連性を調査したところ、欧州では同じ先行バリデーション研究が複数のPASSで再利用されていることがわかった。また、同じアウトカムでも対象患者を絞ってバリデーション研究を再実施しているケースもあった。（図1）

図1：アウトカムバリデーション研究結果の再利用状況



2. 各疾患における国内外比較

脳梗塞・感染症・癌について、日米欧のアウトカムの定義・バリデーション方法・結果をとりまとめ、比較したところ、以下の傾向が見られた。(図2)

- ・日本と海外のアウトカム定義は必ずしも同一ではない。
- ・日本のアウトカムバリデーションが「カルテレビュー」のみに対して、海外では「医師への調査票送付」や「他データとのリンケージ」等、手法が多岐に渡る。

図 2：アウトカム定義とバリデーションの国内外比較（例：脳梗塞）

		日本	イギリス	イギリス	デンマーク	スペイン	米国	
		MID-NET	THIN	CPRD	DNRP	SIDIAP	Sentinel	
情報元 (DOI)		Tanigawa (10.1002/pds.9423)	Ruigomez (10.1002/pds.1919)	Hall (10.1212/01.wnl.000011 0312.36809.7f)	Kranup (10.1159/000102143)	Indacaterol/ Glycopyrronium Bromide	Ammann (10.1097/MD.0000000000 0008440)	Coding Trend Fung
種類		acute ischemic stroke	ischemic stroke /TIA	stroke	ischemic stroke	stroke	acute ischemic stroke	ischemic stroke
定義	概要	傷病名十入院 十検査十薬剤 (Algorithms 8)	傷病名十入院 十フリーテキスト (Read Code)	傷病名 (Read/OXMS Code)	傷病名十入院 (ICD10)	傷病名 (ICD10)	傷病名十入院 (ICD9)	傷病名 (ICD9/ICD10)
	詳細	I63.X/G08 CT/MRI/MRA 抗血栓薬	G63-G67 /F423600	NA	I63.X	I60-I69 /G46/Z86	433.x1/434.xx/436	433.x1/434.xx/436 I63.X
方法	対象	一般集団	一般集団	片頭痛患者	一般集団	COPD患者	IGIV投与者	-
	手順	カルテレビュー (3施設) ① Possible Caseから サンプリング (366例) ② 2人の医師による 真偽判定 ③ 11種類の定義別 でPPV算出 (1施設で感度算出)	医師への調査票 ① アウトカム定義群 からサンプリング (245例) ② 医師に調査票を 送付して真偽判定 (161例の回答) ③ PPV算出	データリンク (Written recordsと) ① アウトカム定義群 からサンプリング (339例) ② Written recordsの 取得(215例) ③ Computerized recordsとの比較で 真偽判定 ④ PPV算出(78/88)	データリンク (コホート研究と) ① アウトカム定義群 をCopenhagen City Heart Study とクロスリンクし、 詳細な情報を入力 (236例) ② 2人の評価者による 真偽判定 ③ PPV算出(32/33)	データリンク (入院情報と) ① アウトカム定義群 を抽出(144例) ② リンクした入院情報 で真偽判定 ③ リンク無しの症例 は医歴記録を マニュアルレビュー ④ PPV算出	カルテレビュー (13のSDD Partner) ① アウトカム定義群 からサンプリング (128例) ② 数名の医師による 真偽判定 ③ 病名ポジション別 (principal/secondary) にPPV算出	2015年10月の ICD9からICD10 への移行を機に、 そのマッピング方法 および移行状況を 解析 (脳卒中の発生率 推移を解析した所 移行前後での差異 はなかった)
結果	PPV	84.3% (Algorithms 8)	85.7%	89%	>97%	75.7%	60% (Principal position)	-
	感度	88.9% (Algorithms 8)	-	-	-	-	-	-

(2) 日本のアウトカムバリデーション手法

日本のアウトカムバリデーションは、論理的なアルゴリズムを作成後、リアルワールドデータからアルゴリズムで症例を抽出し、医師レビューやアブストラクション、疾患DBとの突合等のカルテレビューにより検証し、妥当性の指標となるPPV (Positive Predictive Value) や感度等を算出する。そして、さらに精度向上のために機械学習等も用いたアルゴリズムの修正を必要に応じて実施し、最終的なPPVや感度を算出する。

1. 日本の事例

事例1：確定病名のみアウトカム定義

1型糖尿病の確定診断のみをアウトカム定義とした場合にPPVが52%であったが、2つのインスリン治療法(CSII (Continuous Subcutaneous Insulin Infusion) 用のシリンジポンプとインスリン針)のクレームコードを症例定義に加えることで、PPVが90%を

越えることができた。本事例より、保険傷病名だけの定義では、一般に PPV は低いことが分かる。

事例 2：疾患レジストリの利用

肺がん関連病名の付与をアウトカムとし、真偽判定に院内がん登録を用いたところ、真偽判定労力を削減できた。それにより症例数の制限を緩和でき、初期定義は低い PPV から開始できた。但し、これには該当する疾患がすべて登録されている（悉皆である）ことが必要であり、院内がん登録はその条件を満たしていた。院内がん登録との突合によりアウトカム定義を修正したところ、PPV13.1%・感度 99.8%から PPV83.8%・感度 88.7%と改善できた。本事例より、悉皆性の高い疾患レジストリ等があれば良いゴールドスタンダード(以下、GS)となり、効率よく正確に感度・特異度・PPV・NPV (Negative Predictive Values) を把握できることが分かる。

事例 3：アブストラクション及び疾患レジストリの利用

訓練されたアブストラクターがランダムに抽出された 20 症例で予備のアブストラクションをし、専門医と項目などを修正し全部で 200 症例(各施設)をアブストラクションした。医師レビューとアブストラクションの結果に基づく評価では、PPV に差はなく十分に高かった。疾患登録 DB を GS とした場合は、登録率が低いため PPV が低く出た。本事例より、訓練されたアブストラクターであれば専門医の直接のカルテレビューと判定結果は変わらない。一方、疾患レジストリも悉皆性が低ければ、良い GS にはならないことが分かる。

事例 4：各施設で別々のアブストラクターがアブストラクション

アブストラクターを各協力医療機関でリクルートし、Web 説明会に全員が参加した上で、アブストラクションフォームに従ってアブストラクションを行った。その結果をもとに専門医師 2 名が真偽判定し、過去の専門医レビューの真偽判定結果を GS として比較した。各施設のアブストラクターの経験や職種はばらつきが大きく、検証の結果、一致率は十分ではなかった。アブストラクターを医療施設毎に確保することは、アウトカム検証の精度上でも困難なことが示唆された。

事例 5：All Possible Cases を設定

All Possible Cases (APC) 定義とは、「対象とする事象」を 100%拾えると想定される集団を特定する広義の定義であり、PPV、感度の両方を効率的に算出できる。入院して蛋白分解酵素阻害薬の経静脈的投与による治療を要した急性膵炎または慢性膵炎の急性増悪を対象事象として作成した APC 定義では、PPV51.2%・感度 100%を得た。その後、アウトカム定義を修正検討した結果、PPV71.6%・感度 80.7%というアウトカム定義が作成できた。本事例より、APC 定義を用いて効率的にアウトカム定義の検討をすることで、感度の低下を最小限としつつ、PPV を改善できる場合があることが分かった。

事例 6：偽陽性をゼロにするため、アウトカムが明確なコホートを活用

テネシーの Medicare データで突然死した人のデータを使った研究では、登録されている症例は、全例が突然死であることが確認されているため偽陽性はない。PPV、特異度は 100%となるため、検証の必要がなく、最終的に出てくるリスク比にバイアスがかからない。本事例より、アウトカムが明確なコホートを使って偽陽性をゼロにすることができれば、そのデータが効率的に利用可能となることが分かる。

以上の事例より、以下の事項が示唆された。

- ・ 悉皆性の高い疾患レジストリが利用可能な場合には効率的な検討が可能である。なお、本邦ではそのような疾患レジストリは少なく、他のデータとの突合が制度上制限される場合もあり留意が必要である。
- ・ 「疾患名」のみ等の簡単なアウトカム定義は精度が低いことが多いため、アウトカム定義の検証や改善は重要である。
- ・ 専門医によるカルテレビューは医師の負担が大きいため、アブストラクションの手法や人材をさらに洗練化する必要がある。
- ・ 入力情報の精度向上のため、いずれ電子カルテ等の入力機能の改善が必要である。

(3) 海外のアウトカムバリデーション手法

1. 医療 DB の種類に応じたバリデーション手法

① 統合型 DB

社会保障番号等で地域の各種データがすべて繋がっている。デンマークやカナダに見られる。バリデーション研究数が少ないが、これは全てのデータが繋がっているため、バリデーションの必要性が乏しいことが理由と推測される。

② プライマリーケア型 DB

かかりつけ医（以下、GP）中心のデータで、連携先の専門医や病院における診断名等も入っていることが多い。イギリスやスペイン、オランダに見られる。バリデーションは GP が主体で行われる。

③ 病院 DB

病院中心のデータで、GP のデータは含まれない。日本では MID-NET がこれに該当する。バリデーションは病院が主体で行われる。

④ 診療報酬請求 DB

電子カルテは含まないが、病名等含む診療報酬請求のデータが施設横断で繋がっている。アジアで多く見られ、日本では NDB がこれに該当する。データ発生源に遡ることが難しいため、バリデーションは複数の病院内で行うことが多い。

2. 海外の事例

事例1：オーバービュー

国際薬剤疫学会の雑誌で各国のバリデーション研究が特集されており、その多くはカルテレビューを実施して PPV のみ求めているが、一部はデータリンケージにより PPV だけではなく感度と特異度も求めている。

事例2：英国

英国では、COPD 病名の付いた患者からランダムに選択して GP に質問状を送付し、その回答結果から PPV を計算している。GP から真の COPD 患者全ての情報を得るには膨大な労力がかかるため、この手法では感度の算出は難しい。GS となるレジストリ等とリンケージできる場合は、PPV だけでなく、感度と特異度も算出しやすい。

事例3：米国

米国のメディケア Claims データを活用した研究では、バリデーション研究を行っておらず、代わりに Sentinel の先行研究結果を適用してアルゴリズムを決めていた。同じ Claims データであることから適用できると判断したと考えられるが、異なる DB であることから許容できない場合もあり、論文であればレビューアーやエディターの判断になる。このように外挿が妥当かどうかは、一律の判断が難しい。

事例4：台湾

台湾では、実施する研究と同時に、その病院内でバリデーション研究も併せて実施していることが多い。データのアクセスが容易な自病院で両研究を実施していると考えられる。この場合も、PPV だけ計算されていた。

以上の事例より、以下の事項が示唆された。

- ・米国や台湾は、単・複数の病院の中でバリデーション研究を行い、その結果を発生源が類似する他の DB 研究に外挿している。
- ・日本でも、外挿の限界をよく踏まえた上で、米国や台湾と同様のアプローチを取ることが現実的な選択肢となる可能性がある。

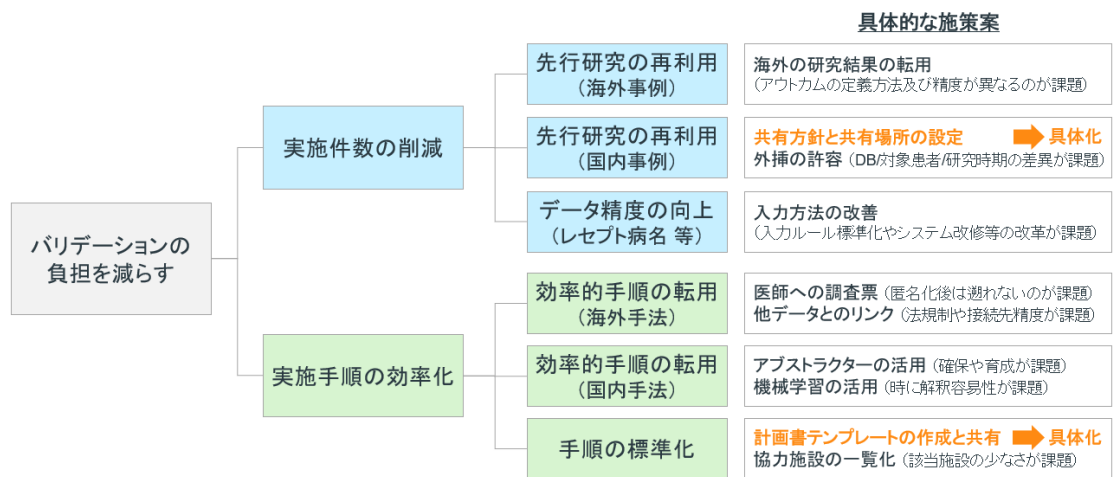
アウトカムバリデーション効率化の施策

(1) 討議であがった施策一覧

「国内外のアウトカム定義とバリデーションの比較」や「日本と海外のアウトカムバリデーション手法の事例」を基にした討議から、アウトカムバリデーション効率化の施策を抽出した。(図 3)

抽出した各施策の具体化を検討し、「先行研究の共有方針と共有場所の設定」と「計画書テンプレートの作成と共有」についてはより具体的な内容に落とし込んだ。

図 3：討議であがったアウトカムバリデーション効率化の施策



1. 具体的な施策の検討

先行研究の再利用 (海外事例)

日本と海外では医療の制度や環境が異なるため、海外の結果をそのまま利用することは難しい。一方で、計画段階におけるアルゴリズムの検討に海外事例を参考にすることが有用な場合がある。

先行研究の再利用 (共有方針と共有場所の設定)

令和3年度のコンソーシアムにおける議論の結果、バリデーション結果の共有と再利用について検討していくことが有用とされたため、その具体策について検討を行った(詳細は(2)1.項で記載)。

先行研究の再利用 (外挿の許容)

結果の外挿にあたっては、DB の特徴、対象とする事象や患者集団、研究時期等における差異などを考慮する必要がある、基準や閾値を一律に設定するのではなく、個々に判断する必要がある。

データ精度の向上

電子カルテの入力ルールやコードの標準化等が必要で、システム改修等も必要となるため、今すぐに解決できる課題ではないが、長期的視点に立ち検討を進めていく必要がある。なお、医療情報学会においてもデータ精度の向上のために検討が進められており、電子カルテメーカーとの連携も含めた検討が始まっている。

効率的手順の転用（海外手法：医師への調査票、他データとのリンケージ）

日本では、現状、匿名加工したデータの原資料に遡ることや匿名加工 DB 同士のデータリンケージは一部の DB を除いて法律上認められていない。そのため、匿名加工前の医療情報を有する同一病院内において、DB とのリンケージが可能な院内がん登録等の疾患レジストリを可能な限り活用していくことが現時点においては現実的である。

効率的手順の転用（国内手法：アブストラクターの活用）

アブストラクターによる支援は、真偽判定する医師の負担軽減にもつながる可能性があるが、アブストラクターに対する事前の十分なトレーニング等が必要である。なお、各病院がアブストラクターを確保してトレーニングを実施することは困難と考えられることから、現実的にはトレーニングを受けた外部アブストラクターを活用することが考えられるが、その費用に対する予算確保等が必要となるため、アブストラクターの要件等を明確化し、費用の予見可能性を高めることが求められる。

効率的手順の転用（国内手法：機械学習の活用）

機械学習の活用は、カルテレビューにおける真偽判定プロセスとしては追加的な要素であるが、アルゴリズムの精度向上において有用な手法であると考えられる。

手順の標準化（計画書テンプレートの作成と共有）

実施計画書の記載項目について共通認識をもつことは、一定程度質が担保されたアウトカムバリデーションの効率的な実施を促進することにつながるため、PMDA にてアウトカムバリデーション計画書のテンプレートを作成し共有した（詳細は(2)2.項で記載）。

手順の標準化（協力施設の一覧化）

協力施設を一覧化することは、アウトカム定義のバリデーションが実施可能な施設を可視化する観点からは有用であるが、一方で、バリデーション実施施設における負荷の集中や実施可能施設の限定化などのおそれもあるため、現時点では一覧化に向けた対応ではなく、実施可能施設の増加に向けて検討を進めていく必要がある。

(2) 具体的な施策案

1. 先行研究の再利用（共有方針と共有場所の設定）

海外の規制当局サイトにおける共有方法を調査し、日本の規制当局サイトにおける共有方法を具体的に検討した。

海外の規制当局サイトでの共有方法調査

脳梗塞を事例に、海外の規制当局主導の WEB サイトにおける先行研究の共有方法を簡易的に調査した。（図 4）

先行事例を一つにまとめた情報は、米国では、Literature Reviews の結果をアウトカム別に研究一覧として掲載していたが、欧州では同様の情報掲載はなかった。また、個別事例の掲載に関しては、米国、欧州とも検索された各研究の補足資料として格納されていることが多く、比較的サイト内の深い階層に格納されていた。

なお米国では、検索用にアウトカムバリデーション研究の種類やアウトカム種別がタグ付して整理されているため、事例検索が容易な仕組みとなっていた。

図 4：海外の規制当局サイト上での先行研究の共有方法

		米国 (FDA)	欧州 (EMA)
		Sentinel	RMP (1/2/3)
		Drug Assessments Sentinel Initiative	EU PAS Register (encepp.eu)
事例まとめ (一覧表)	掲載場所	Health Outcomes of Interest	-
	掲載内容	Systematic review結果 - Sentinel以外のDBのバリデーション研究 - アウトカム別に研究を一覧でまとめている - 研究内容/結果および文献情報等を要約	-
個別事例	掲載場所	各PASSの要約サイト内（掲載ない場合も有り）	各PASSの要約サイト内（掲載ない場合も有り）
	検索方法	① 「Search Health Outcomes of Interest」で該当PASS/Validationをフィルタリング（TitleとOutcomesが検索対象でタグ付き） ② 検索先の「Additional Information」タブ内の格納を確認	① 「Search Study」で該当PASS/Validationを検索（Titleが検索対象） ② 検索先の「Document」タブ内の格納を確認（格納ない場合はPASS Protocolの中身を確認）

日本の規制当局サイトでの共有方針

製販後 DB 調査の計画の参考とするため、「製造販売後データベース調査で用いるアウトカムバリデーション実施に関する基本的考え方」に準拠している研究を対象として医薬品医療機器総合機構（以下、PMDA）の WEB サイト上で共有していく方針とした。研究論文の著者等が許諾したものに限り掲載することとし（図5）、具体的には対象とするアウトカム毎に、各論文のアウトカム定義・研究方法・結果・リンク先を掲載するイメージである。

図 5：日本の規制当局サイト上での掲載内容

掲載の場所	PMDAのWEBサイト上
掲載の目的	製造販売後データベース調査を計画する際に参考となるアウトカムバリデーション研究を共有する
掲載対象の要件	<p>「製造販売後データベース調査で用いるアウトカムバリデーション実施に関する基本的考え方*」に準拠している研究で、著者等から掲載許可を得たもの *薬機レギ長発第0731002号及び薬機審長発第0731002号</p> <ul style="list-style-type: none"> - 陽性的中率だけでなく感度も算出 - 複数施設での実施 - ケースの真偽判定に専門医が関与

なお、「製造販売後データベース調査で用いるアウトカムバリデーション実施に関する基本的考え方」の要点は以下の通りである。

- (1) 製販後 DB 調査においては、対照群に対する相対リスクを適切に算出する観点から、アウトカム定義の妥当性評価において PPV は重要な指標である
- (2) アウトカム定義により特定された集団が、対象となるアウトカム全体のうち、どの程度を網羅できているかを把握する観点から、原則として、感度についても検討すべきである
- (3) バリデーション実施施設は、バリデーション結果の一般化可能性を高めるため、利用予定の医療情報 DB で代表性を有する複数施設を選定することが必要である
- (4) カルテ内容を精査して、真のケースを判定する場合には、原則として、1 症例につき 2 名以上の専門医が独立して判定するなど、一定程度専門医が関与して検討すべきである

2. 手順の標準化（計画書テンプレートの策定と共有）

PMDA ガイダンス

アウトカムバリデーション研究を促進するために、「アウトカム定義のバリデーション研究実施計画書作成ポイント（案）」を PMDA で作成した。（図 6）

本資料は、一般的に記載すべき事項を網羅した計画書の作成において参考となり得る資料であり、薬剤疫学調査、製造販売後データベース調査等で用いるアウトカム定義のバリデーションに関する検討を促進し、本邦におけるデータベース調査の質の向上に繋がることを期待できる。

図 6：計画書テンプレート（案）

名称	アウトカム定義のバリデーション研究実施計画書 作成ポイント(案)	
目的と位置づけ	アウトカム定義のバリデーションに関する検討を促進し、 その質の向上のために、一般的に記載すべき事項を参考 としてまとめたもの	
資料の構成	<ul style="list-style-type: none"> 1. 研究計画書更新履歴 2. 研究計画書の概要 3. 検討の対象とする事象 4. 用語の定義 5. 略語 6. 研究の工程 7. 研究実施体制 8. 研究の目的 9. 研究に用いる医療情報 	<ul style="list-style-type: none"> 10. 研究で用いるデータ項目 11. 研究の方法 12. 個人情報保護・倫理 13. 資金源・利益相反 14. 引用文献 15. 付録