

医薬品規制調和国際会議

ICH 調和ガイドライン

医薬品の安全性評価においてリアルワールドデータを活用
する薬剤疫学調査の計画、デザイン、解析に関する
一般原則

M14

草案

2024年5月21日承認

現在パブリックコメント（意見公募手続）中

ICH プロセスのステップ 2 では、ICH 専門家作業部会で合意された合意文書又はガイドラインの草案は、各国又は各地域の手順に従い、内外の意見収集のため、ICH 総会によって ICH 各地域の規制当局に送付される。

M14
文書の履歴

ガイドライン	履歴	日付
M14	Step2 の ICH 運営委員会による承認及びパブリックコメント（意見公募手続）のための公表。	2024 年 5 月 21 日

法律上の注意

本ガイドラインは著作権により保護されており、ICH のロゴを除き、常に ICH の著作権が認められていることを条件に、パブリック・ライセンスの下で使用、複製、他の著作物への取り込み、改変、修正、翻訳、配布することができます。文書の改変、修正、翻訳が行われた場合は、原文に変更が加えられたこと又は原文に基づくものであることを明確に表示、区分又はその他の方法で識別するための合理的な措置を講じなければなりません。原文書の改変、修正、翻訳が ICH によって承認又は後援されているかのような印象を与えることは避けなければなりません。

文書は、いかなる保証もなく「現状のまま」提供されます。いかなる場合においても、ICH 又は原文書の作成者は、本ガイドラインの使用から生じるいかなる請求、損害、又はその他の責任についても責任を負いません。

上記の許可は、第三者から提供されたコンテンツには適用されません。したがって、著作権が第三者に帰属する文書については、その著作権者から複製許諾を得なければなりません。

医薬品の安全性評価においてリアルワールドデータを活用する 薬剤疫学調査の計画、デザイン、解析に関する一般原則

M14

ICH 調和ガイドライン

目次

1 はじめに	1
1.1 目的	1
1.2 背景	1
1.3 適用範囲	2
1.4 医薬品の安全性評価以外の目的で実施する調査	3
2 一般原則	3
3 リアルワールドデータを用いて適切なエビデンスを生成するためのフレームワーク	4
4 調査の初期デザインと実施可能性	6
4.1 リサーチクエスチョン	6
4.2 フィージビリティ評価	7
5 調査実施計画書の作成	10
5.1 調査デザイン	11
5.2 データソース	12
5.2.1 注目する安全性の検討課題への対処におけるデータソースの適切性	13
5.2.2 主なデータソースの特徴	14
5.2.3 データの標準化	20
5.2.4 欠損データ	20
5.2.5 データの品質	21
5.2.6 調査実施計画書におけるデータ収集及びデータソースに関する項目	21
5.3 対象集団／調査の対象集団	22
5.4 曝露、アウトカム、共変量	22
5.4.1 曝露	25
5.4.2 アウトカム	26
5.4.3 共変量	28
5.5 バイアスと交絡	30
5.5.1 選択バイアス	31
5.5.2 情報バイアス	31
5.5.3 不死時間バイアス	31
5.5.4 交絡	32
5.6 データのバリデーション	32

6 データマネジメント	33
6.1 データ保有者	34
6.2 調査実施者	35
7 解析	35
7.1 統計解析	36
7.1.1 主要解析	36
7.1.2 欠損データ	37
7.1.3 感度解析	37
8. 報告及び提出	38
8.1 有害事象、副作用及び製品品質に関する苦情の報告	38
8.2 規制当局に提出する調査の文書書式と内容	38
9 調査の資料並びに結果の公表及び伝達	38
10 調査の文書及び記録の保管	39
11 特定の集団における検討事項	40
11.1 妊娠に関する調査	40
12 用語集	42
13 関連する規制ガイドライン	47
14 規制当局以外による関連するガイドライン	49
15 参考文献	49

1 1はじめに

2 1.1 目的

3 本ガイドラインの目的は、医薬品（薬剤、ワクチン及びその他の生物学的製剤を
4 含む。以下同様。）の安全性評価のために目的に適したデータを活用する観察（非介
5 入）方式の薬剤疫学手法を用いた調査（以下、「調査」という）の計画、デザイン及
6 び解析に関する一般原則について、国際的に推奨する基準を示し、国際調和を促進
7 することである。

8 本ガイドラインでは、調査の実施計画書及び報告書の作成と規制当局による評価
9 を合理化するために、これら調査の実施に関する推奨事項及び高水準のベストプラ
10 クティスの概要を述べる。これら推奨事項及びベストプラクティスは、調査の実施
11 計画書及び／又は結果について、規制当局間での受け入れ可能性の向上及び調査結
12 果に応じた意思決定を支援することを目指している。用語集では、本ガイドライン
13 の目的の範囲でいくつかの用語を定義している。用語集で定義された用語について
14 は、初回使用時に**太字のイタリック体**で示した。

15 1.2 背景

16 薬剤疫学調査は、製造販売承認を受けた**医薬品 (medicine)** の市販後安全性の評
17 価を裏付けるデータ及びエビデンスの情報源として長く用いられてきた。

18 シグナル^{*}は様々なデータソースから生じる可能性がある。これは、例えば、製品
19 の品質、非臨床データ及び臨床データ（医薬品安全性監視及び薬剤疫学的データを
20 含む）等、医薬品の使用とその結果に関するすべての臨床的及び科学的情報を含む
21 可能性がある。薬剤疫学調査は、安全性の懸念又はシグナルの検出、特徴の把握及
22 び評価の重要な要素であり、記述的又は推論的に実施する場合がある。

23

24 （※ 医薬品の安全性に関するシグナル。「定期的ベネフィット・リスク評価報告
25 （PBRER）について」（平成 25 年 5 月 17 日付薬食審査発第 0517 第 1 号、厚生労働
26 省医薬食品局審査管理課長通知）参照。）

27

28 規制目的で利用する頑健なエビデンスの生成は、**データの信頼性 (data**
29 **reliability)** 及び**データの適合性 (data relevance)** に加え、そのデータを解析するた
30 めの適切な薬剤疫学的手法の利用により成り立っている。規制上の意思決定のため
31 の薬剤疫学調査の利用は世界的に増加しており、規制当局や専門家団体によって複
32 数のガイドライン及びベストプラクティスが作成されてきた。多くの国や地域で
33 は、主に医薬品の安全性評価を目的とする薬剤疫学調査の計画やデザインについて
34 の一般原則に関するガイドラインが発出されている。さらに、The Sentinel Innovation
35 Center の PRINCIPLED フレームワーク、ISPE/ISPOR の HARmonized Protocol
36 Template to Enhance Reproducibility (HARPER) イニシアチブ等、調査デザイン及び
37 実施についての枠組みが非政府グループによって作成されており、これらには本ガ
38 イドラインの範囲を超えた追加の詳細が記載されている[1,5]。

39 1.3 適用範囲

40 **リアルワールドデータ (RWD, real-world data)** を構成する要素については地域に
41 より若干の差異があると認識しているが、本ガイドラインでは、医薬品の市販後安
42 全性を評価する目的で実施される RWD を活用した調査に関する推奨事項を示す。注
43 目するリサーチクエスチョンへの回答にあたって、既存の RWD の活用だけでは不十
44 分な場合、調査実施者は調査の実施を目的として追加のデータを収集することがあ
45 る。本ガイドラインでは、特定の調査の実施を目的として収集されたデータを**一次**
46 **データ収集 (Primary Data Collection)** という。一次データ収集は RWD を用いた観
47 察方式の調査に有用な場合があるため、本ガイドラインでは一次データ収集に関す
48 る検討事項も合わせて述べる。

49 臨床試験又は薬剤疫学調査のどちらが最も適切なアプローチであるかについての
50 考え方を提供することは、本ガイドラインの範囲を超えるものである。また、本ガ
51 イドラインは、薬剤疫学的手法に関する包括的な知識を提供することを意図したも
52 のではない。むしろ、本ガイドラインは、薬剤疫学調査のデザイン、計画及び実施
53 に関する規制当局のガイダンスを調和させ、規制当局による評価を容易にすること
54 を意図している。また、規制当局のガイダンスで取り上げられていない範囲につい
55 ては、その他の資料にとりまとめられたベストプラクティスのガイダンスを適宜考
56 慮することもできる (14 章「規制以外のガイドライン」を参照)。

57 以下の臨床試験、研究又は医薬品安全性監視の目的で実施するその他の活動は本
58 ガイドラインの適用範囲外である。:

- 59 • 国内又は海外のデータベース（例：国レベルの医薬品安全性監視システム）か
60 ら得られた自発報告を用いた医薬品安全性監視における調査。
- 61 • 無作為化比較試験、プラグマティック試験、実施計画書に従って治療方法が定
62 義された単群の臨床試験、外部比較対照を用いた試験等の治療割り付けを伴う
63 試験。
- 64 • **患者体験データ (patient experience data)** を収集し解析する調査。

65 患者体験データの収集は、注目すべき事象、視点、ニーズ及び優先順位等の観点
66 に関する情報を提供する上で、市販後安全性評価において重要な要素となり得る。
67 患者体験データの収集に関する詳細なガイダンスは本ガイドラインの範囲外である
68 が、いくつかの規制ガイダンスが作成されている（第13節「参照する規制ガイドラ
69 イン」を参照）。調査に患者体験データが含まれる場合、調査実施者は追加の情報と
70 して、これら関連する既存のガイダンスを参考にすることができる。

71 ファーマコゲノミクス、人工知能（AI）及びRWDの使用に関連するその他の新た
72 な技術については、それら技術の発展性を考慮し、これらのトピックについては本
73 ガイドラインで取り上げない。

74 1.4 医薬品の安全性評価以外の目的で実施する調査

75 本ガイドラインに示す原則は、医薬品の使用状況、有効性に関する検討等、医薬
76 品の安全性の評価以外の目的で実施される市販後の薬剤疫学手法を用いた調査にも
77 適用できる推奨事項を含んでいる。それらの調査に RWD の要素が含まれる場合に
78 は、本ガイドラインに示された基本原則を参考にしてもよい。

79 2 一般原則

80 医薬品の安全性プロファイルは、非臨床試験から製造販売承認後のライフサイク
81 ルを通じて蓄積される知識体系の発展を反映したものである。製造販売承認後にお
82 ける医薬品安全性評価のための薬剤疫学調査は、実臨床で使用される医薬品のベネ
83 フィット・リスクプロファイルのより明確な全体像を示すために、他の情報源を補

84 完するものである。

85 本ガイドラインでは検討の工程を段階的に記述しているが、調査デザインの検討
86 及びデータソースの選択における様々な工程は段階的に繰り返し行われるものであ
87 る。検討の流れはリサーチクエスチョンの明確化から始まり、調査の対象集団、**曝**
88 **露 (exposure)**、アウトカム及び共変量の特定、調査の実施可能性の評価（以下、
89 「フィージビリティ評価」）を行うための最小限のデータ要件の特定、対象集団に対
90 するデータソースの代表性の評価、並びに潜在的な**バイアス (bias)** 及び**交絡**
91 **(confounding)** の原因を検討するための体系的な検討を行う。適切なデータソース
92 及び／又はデータ収集方法が特定された後、調査の妥当性に対処する手法も含め
93 た、調査デザインをさらに洗練する検討を行う。目的への適合性の評価に関する 3
94 章「リアルワールドデータを用いて適切なエビデンスを生成するためのフレームワ
95 ーク」には、これらの活動の総合的な検討の流れを示す。この検討の手順を通じ
96 て、曝露、アウトカム及び交絡因子の定義、並びに解析、データマネジメント、調
97 査の実施、報告、提出、結果の公表、その他の重要な決定事項についての基本とな
98 る理論的根拠及び正当化する根拠は文書化されるべきである。また、調査における
99 運用上のすべての事項は、明瞭かつ透明性があるべきである。

100 本ガイドラインにおいて「調査実施者」とは、規制当局、調査の依頼者、医薬品
101 開発業務受託機関、学術団体又はその他の調査の計画及び実施に責任を有する者を
102 いう。規制当局に提出される製造販売後の医薬品安全性評価のための薬剤疫学調査
103 に関するすべての事項については、製造販売承認申請を行う者及び医薬品製造販売
104 業者が最終的な責任を負うものとする。

105 3 リアルワールドデータを用いて適切なエビデンスを生成するためのフレームワー 106 ク

107 規制上の意思決定を支援するために提出される、調査から生成されたエビデンス
108 の強さは、基礎となるデータの適合性及びデータの信頼性に加えて、調査デザイン
109 及び方法論に依存する。適切なエビデンスを生成するためのフレームワーク（図
110 1）では、まずリサーチクエスチョンが設定されるべきであり、次にそのクエスチョ
111 ンに対処するために最も適切な調査デザイン及びデータソースが決定される（2）。

112 特定のデータソースにおいては調査デザインの選択肢が制限され、導き出すことが
113 できる推論を制限する可能性があることから、調査実施者は、特定のデータソース
114 にのみ適合する調査を策定することは避けるべきである。一般に、得られるエビデ
115 ンスがリサーチクエスチョンに答えるのに十分であるかどうかを判断するためのフ
116 レームワークには、次の (1) ~ (3) の総合的な評価を含むべきである。

117 (1) データの適合性及びデータの信頼性

118 (2) 調査デザイン及び解析手法の適切性

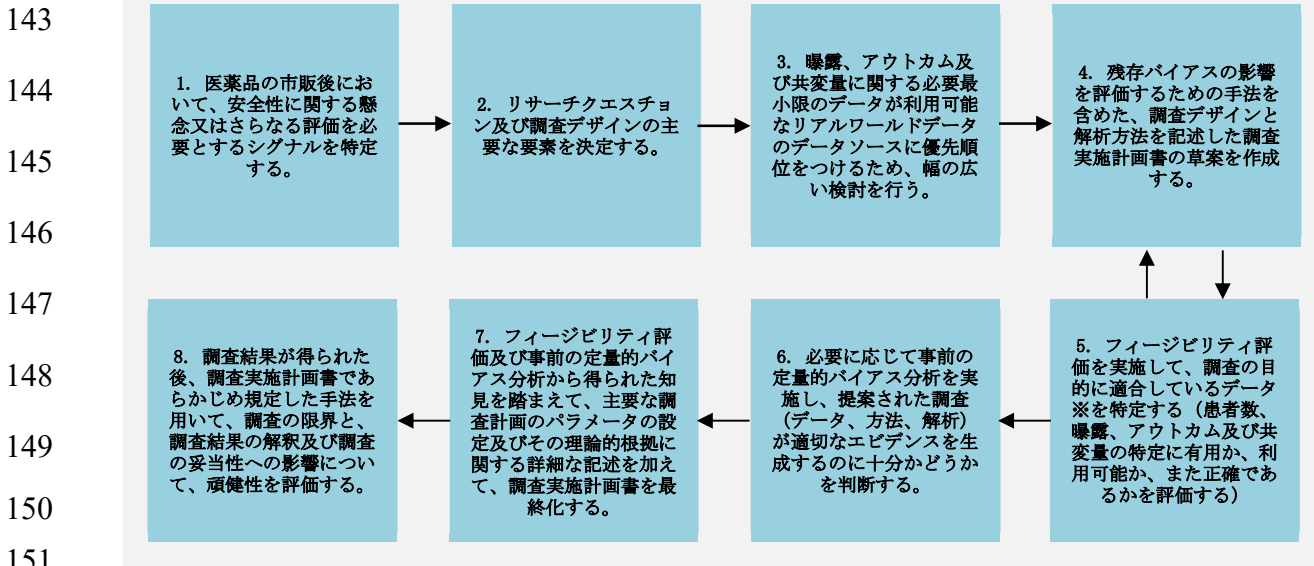
119 (3) 調査の限界に関する定性的／定量的な頑健性の評価、並びに得られたエビデ
120 ス及び結果の解釈についての最終的な妥当性及び信頼性への影響

121 調査から得られたエビデンスが十分かどうかの総合的な評価は、調査実施計画書の
122 検討段階における実施可能性の評価（例：理論的な懸念が及ぼす影響の考察、バイ
123 アスの原因となり得る因子やそれらが調査の妥当性に及ぼし得る影響）及び調査計
124 画書に事前に規定された感度解析を実施する調査実施後の段階の両方で検討される
125 べきである。**定量的バイアス評価（解析）(quantitative bias assessments**
126 **(analyses))** は、フィージビリティ評価のために事前に行う場合、調査結果の解釈
127 を容易にするために行う場合又はその両方の目的で用いることができる。フレーム
128 ワークに含まれる 3 つの要素を考慮することで、調査実施計画書に従って実施され
129 る調査によって、注目する規制上の課題に対処しうる適切なエビデンスが得られる
130 可能性があるか判断ができるようになる。他のデータソース（例：ウェブサイト、
131 ブログ、ソーシャルメディア、チャットルーム）から取得したユーザ投稿型の健康
132 に関するデータを用いた調査では、注目する規制上の課題への対処には十分ではな
133 い場合があるが、仮説の生成及び調査結果の位置づけを明確化するための参考情報
134 が得られることがある。

135 図 1 には、調査計画の検討の工程を直線的に示すが、規制上の課題に対処しうる
136 十分なエビデンスを得られるよう、調査計画の策定における検討及び評価は反復的
137 に行うべきである[2]。調査実施者は、調査計画の策定の早期の段階において、調査
138 の特性について規制当局と話し合うことが推奨される。次章では、調査計画の妥当
139 性の評価を可能とするために、調査実施計画書に必要な要素について概説する。

140 図 1：医薬品安全性に関する規制上の問題に対処するために、目的に適合したり
 141 アルワールドデータを使用して適切なエビデンスを生成するためのフレームワーク

142



152 ※ 電子診療録（以下、EHR）レセプトデータ、一次データ収集又は他の種類の RWD の単独又は組み合わせ。

153

154 4 調査の初期デザインと実施可能性

155 4.1 リサーチクエスチョン

156 リサーチクエスチョンとは、調査の目的及び事前に規定した検討すべき仮説を簡
 157 潔に述べたもので、将来の調査のための仮説を生み出すことを調査の目的とするこ
 158 ともある。リサーチクエスチョンは、調査の対象集団、介入（非介入の調査の場合
 159 は曝露）、比較対照、アウトカム及び調査の対象期間（PICOT）を含むテンプレート
 160 を用いて策定することができる。非介入の調査の場合、「介入」は曝露と同義である
 161 とみなすことができる。リサーチクエスチョンは、不足している知見、先行する調
 162 査、試験又は研究の長所及び短所、期待される効果の大きさ、及び重要な交絡因子
 163 を特定し理解するために、文献を検討した後に策定されるべきである。リサーチク
 164 エスチョンの策定においては、調査の具体的な目的によってリサーチクエスチョン
 165 がどのように解決されるかを示す明瞭な理論的根拠を提示する必要がある。調査実
 166 施者は、調査実施計画書において、調査デザイン及び必要とされる／利用可能なデ

167 ータの種類について、その判断を支持するよう記載すべきである。リサーチクエス
168 チョンを慎重に設定することで、フィージビリティ評価から得られた情報を踏まえ
169 て取り組むべき未知の課題を明確にでき、この情報によってリサーチクエスチョン
170 がさらに洗練され、調査計画の発展が促進される可能性がある。また、調査実施者
171 は、リサーチクエスチョンの検討を開始し、調査デザインの草案の作成及び実施可
172 能性の評価を実施する際には、**標的試験 (target trial)** アプローチ又は **estimands** フ
173 レームワーク等の、医薬品のリスクの推定や調査デザインの検討に関する原則的な
174 枠組みに沿って検討することもできる。

175 4.2 フィージビリティ評価

176 フィージビリティ評価は、各治療群のアウトカムを評価することなく、特定のリ
177 サーチクエスチョンに取り組むための目的に適合したデータを特定するとともに、
178 想定される調査の統計的精度に関する情報を取得するための体系的な工程である。
179 フィージビリティ評価を実施する際の主要な目標は、評価対象の医薬品に関連する
180 アウトカムを評価することなく、リサーチクエスチョンの検討における各データソ
181 ースの信頼性及び適合性を比較し説明することである。各データソースの潜在的な
182 利点と制限に関する詳細は、5章「調査実施計画書の作成」に記載する。

183 フィージビリティ評価は、少なくとも2つのフェーズで構成されるべきである：

- 184 • データが利用可能で調査への利用が適切であると期待できるかを判断し、デー
185 タソースの選択肢を絞り込む初期段階での検討
- 186 • 絞り込まれた各データソースについて、詳細な実施可能性の検討

187 リサーチクエスチョン及びデザインの構成要素を決定後、調査実施者は、リサー
188 チクエスチョンに対するデザインの重要な要素に対処するために必要となる最低限
189 の基準を定めるべきである。この作業には、RWD ソースの特性及び臨床状況を理解
190 することが必要になる。検討すべきデザインの構成要素には以下が含まれる。

- 191 • 調査の対象集団、曝露、比較対照、アウトカム及び共変量を把握し定義するた
192 めに必要なデータ
- 193 • アウトカムを観察するための最小の追跡期間

- 194 • 目標とするサンプルサイズ／イベント発生率及び期待される調査の精度
195 • 関心のある地理的地域
196 • 可能な場合には、診断方法、優先的に使用される医薬品、処方されうる医薬品
197 の範囲又は処方業務を含む医療システムに関する情報。

198 非介入の調査デザインを計画する初期段階においては、患者レベルのデータ又は
199 解析データセットへのアクセスに関する必要性を明確にすべきである。調査の依頼
200 者は、規制当局への申請において提出を求められる可能性がある患者レベルのデー
201 タ又は解析データセットについて、規制当局によるアクセスが可能となるよう必要
202 な第三者契約を取得すべきである。

203 フィージビリティ評価に関連する他の重要な要素には次のものが含まれる。:

- 204 • 特に希少疾患の場合、適切な診断コードが使用可能か
205 • アウトカム又は曝露の情報をバリデーションするために、臨床検査による診断
206 の確認及び／又は診療録へのアクセスが必要か
207 • コーディングアルゴリズムの妥当性を示す根拠が存在するか否か

208 リサーチクエスチョンによっては、データソース内の記録を補完するために追加
209 情報を収集できる可能性があるか、データソースを他の種類のデータ（例：バイタ
210 ルサインのデータ、がんのレジストリ、ワクチンのレジストリ）との紐づけができ
211 る可能性があるか等、他の重要な基準を規定することが適切な場合もある。これに
212 より、初期段階での検討時に、調査実施者がリサーチクエスチョンに答えるために
213 重要であると定めた基準を満たす可能性が最も高いデータソースを特定することが
214 できる。場合によっては、公表された情報、データソースの説明、メタデータのカタ
215 ログ、データソースから入手可能な単純な記述統計量等を元に、調査実施者は、
216 初期段階での検討を完了することもできる。

217 調査への活用の候補として利用可能なデータソースを管理可能な数に絞り込んだ
218 後、調査の詳細な実施可能性の検討を実施すべきである。場合によっては、初期段
219 階での検討の際に目的に適合したデータが特定されることもある。その場合、当該
220 データベースに対して、詳細な実施可能性の検討を行うことができる。詳細な実施
221 可能性の検討において、調査実施者は、デザインの重要な要素として必要なデータ

222 が入手可能であること及び調査において最低限必要となるデザイン上の要素について、
223 特定のデータソースにおける妥当性及び完全性に関する十分なエビデンスがある
224 ことを確認する。データソースを選択する際には、データの最新性、データの更
225 新頻度及び曝露からアウトカム発現に至る追跡の完全性について検討すべきであ
226 る。さらに、調査実施中に生成されたデータファイルを規制当局に提出することが
227 可能か、明確にする必要がある。これら検討が必要な事項に対処するために必要な
228 時間と調査結果の取得が求められている緊急性はトレードオフの関係にあることか
229 ら、データに関する事前の知識がデータソースの選択におけるその他の要因となる
230 場合がある。

231 特定されたデータソースを用いたフィジビリティをさらに検討するため、デー
232 タの欠損による潜在的な影響を評価する解析を実施する場合がある。例えば、ホル
233 モン避妊薬と血栓塞栓症との関連性を評価する調査において、喫煙状態に関する情
234 報が欠損していることによる影響を評価するために、喫煙と曝露及びアウトカムと
235 の関連性を文献調査によって明らかにしたうえで、**定量的バイアス解析**
236 (**Quantitative Bias Analysis**) を用いて、期待される効果推定値の範囲に関する調査
237 の妥当性への影響を評価することができる。絞り込まれた各データソースについて
238 詳細な実施可能性を検討するこの工程では様々な情報を利用することから、**データ**
239 **保有者 (data holder)** に特定の情報（例：曝露された患者数、アウトカムの発現率
240 を使用してサンプルサイズを計算するための問合せ、共変量の利用可能性、データ
241 の妥当性を確認するためのその他の問合せ）を要求することも有用となる。

242 詳細な実施可能性の検討が完了した後、データソースを比較し、調査に用いるデー
243 タソースを選択する。これらの工程のいずれかで、特定のデータソースがリサー
244 チクエスチョンに対処するために適していないことが明らかとなる場合がある。こ
245 の場合、調査実施者は一次データ収集の実施可能性を検討することができる。一次
246 データ収集の実施可能性を検討する際には、通常、リサーチクエスチョンに対して
247 意味のある回答を得る上で十分な数の患者を登録可能かどうか及び追跡調査を適切
248 な期間に渡って実施できるかどうかを判断するため、患者集団に関する情報等を医
249 師又は施設に対して問い合わせる。一次データ収集を伴う調査が提案される場合
250 は、調査実施者は調査の立ち上げにかかる時間（施設の選定、倫理的承認の取得、

251 参加者の登録と追跡、結果の報告等) 及びその実施時期が許容可能かどうかを考慮
252 すべきである。

253 さらに、適切な比較対照 (又は調査実施期間) の特定は、調査デザインの重要な
254 部分であり、フィージビリティ評価における重要な検討事項である。健康保険又は
255 医療保険の方針が曝露及び比較対照における疾患重症度の観測度合いに及ぼしえる
256 影響は検討されなければならない、同時比較対照データの利用可能性も検討すべきで
257 ある。ただし、一部の状況 (希少疾患集団の調査等) では、ヒストリカルコントロ
258 ール/従来の**標準治療 (standard of care)** との比較が考慮されることがある。規制
259 当局のガイダンスにて、適切な対照の特徴に関する追加情報を提供している。

260 フィージビリティ評価は、調査実施計画書におけるデザインを判断するための位
261 置づけとして実施される。規制当局と協議する場合又は規制当局に要求された場
262 合、フィージビリティ評価の報告書を単独の文書又は調査実施計画書の別添文書と
263 して提出することがある。当該報告書には、調査計画時に検討されたデータソース
264 に関する情報として、各データソースにおける実施可能性の評価又は探索的な解析
265 の結果を含めて記載すべきである。また、調査の依頼者は、適合性のあるデータソ
266 ースを調査に用いる又は除外することの正当性を説明すべきである。

267 フィージビリティ評価の実施及び報告に関する最終的な手続きは、各地域で適用
268 される規制要件に従うべきである。フィージビリティ評価を実施するための詳細な
269 フレームワーク、テンプレート及びチェックリストは、学術誌に報告されたものを
270 利用することもできる。

271 5 調査実施計画書の作成

272 医薬品安全性評価のための薬剤疫学調査の計画立案及び実施にあたっては、通
273 常、対象分野の専門家の参加が必要である。調査を適切に実施するためには、適切
274 な専門知識を有する経験豊富で、学際的な人員体制が不可欠であり、調査実施計画
275 書には人員体制に含む各担当者の専門知識及び資格に関する記載を含めるべきであ
276 る。人員体制に含める担当者には、以下を含む様々な分野の重要な知見を有するこ
277 とが期待される：

278 • 疾患の症状、因果関係の経路及び現在の実臨床を理解するための適切な医学的

- 279 専門知識に基づく曝露、アウトカム及び共変量の定義の開発
- 280 • 既存の健康・医療データベースにおけるデータ収集の方法及び収集目的に基づ
- 281 くデータベース独自の特徴の理解
- 282 • 疾患領域の診療報酬請求及びコード化の実践経験
- 283 • 一次データ収集の特徴
- 284 • 健康・医療データを利用する際に生じる潜在的なデータプライバシー及び情報
- 285 セキュリティの懸念への対処
- 286 また、データ収集の完全性、曝露、アウトカム及び共変量の評価におけるバイア
- 287 ス、データソースの多様性、データの経時的変化の影響、データへのアクセスの管
- 288 理及び条件、並びにデータベースがカバーする国又は地域の医療システムについて
- 289 も、調査に用いるデータソースの選択に影響を及ぼす可能性がある重要な要素であ
- 290 り、実施計画書で取り扱う必要がある。

291 5.1 調査デザイン

292 医薬品安全性評価のための薬剤疫学調査は通常、対象集団における有害事象の発

293 生率を推定し、医薬品の曝露との関連性を評価することを目的とする。

294 このうち観察（非介入）型の調査では、コホートデザイン、症例対照デザイン、

295 自己対照デザインを含むいくつかの調査デザインが一般的に用いられる。最適な調

296 査デザインの選択は、注目するリサーチクエスチョン及び医薬品の曝露と注目する

297 アウトカムとの間に想定される関係について何が知られているか（例：急性又は遅

298 発性のアウトカム、生物学的な**確からしさ** (*plausibility*)) 等の複数の要因に依存す

299 る。

300 （反事実の経験を反映するようにデザインされた）適切な比較対照集団を特定す

301 ることは、調査デザインの重要な要素である。比較対照の例としては、他の医薬品

302 の使用者、医薬品の非使用者、ヒストリカルコントロール又は自己対照デザインに

303 おける患者自身が挙げられる。対照薬の選択の際に考慮すべき事項には、ある疾患

304 における特徴的な適応、禁忌、疾患の重症度、併存疾患、治療順序等が考えられ

305 る。また、適応の交絡に関連する問題を軽減し、曝露集団と比較対照集団の比較可

306 能性を最大化し評価することが重要である。

307 調査実施者は、調査実施計画書及び最終報告書において、特定の調査デザインを
308 選択した理論的根拠を論じるべきである。また、調査実施者は、解析計画及び組入
309 れ期間 (inclusion period)、ルックバック期間 (look back period)、追跡期間 (follow-
310 up period)、全体の調査期間 (overall study period) 等の時間的要素を明確にするため
311 に、図式化した表現 (調査デザインダイアグラム等) を作成することも検討すべき
312 である。調査デザインの詳細を可視化することで、調査デザインを明確にするとと
313 もに、幅広い意思決定者との意思疎通に役立つ[3]。規制当局における意思決定のた
314 めの適切なエビデンスを提供し得ることを確実にするために、調査計画の立案の早
315 い段階において、調査デザインについて規制当局と協議されるべきである。

316 フィージビリティ評価における初期段階での検討の後に、調査デザイン、解析、
317 実施及び報告に関するすべての重要な要素は特定されるべきである。**データの二次
318 利用 (Secondary Use of Data)** を行う調査においては、選択したデータソースから、
319 それぞれの要素を特定した方法について、適用可能なバリデーションを含めて、調
320 査実施計画書及び最終調査報告書において述べるべきである。

321 5.2 データソース

322 規制当局における意思決定の支援にデータソースを使用する前に、調査の依頼者
323 は、データの適合性及びデータの信頼性を評価することで、そのデータが目的に適
324 合しているかを検討すべきである。本ガイドラインにおいて、適合性という用語
325 は、データの重要な要素 (患者特性、曝露、アウトカム) の利用可能性、及び調査
326 において代表性のある十分な数の患者 (対象集団) の利用可能性を含む。また、信
327 頼性という用語は、**データの正確性 (data accuracy)**、**データの完全性 (data
328 completeness)**、**データの真正性 (data provenance)** 及び**データの追跡性 (data
329 traceability)** を含む。調査実施計画書では、これらデータの主要な特徴についての
330 記述と考察を記載すべきである。

331 薬剤疫学調査においては、調査デザイン及び結果の解釈に影響を与える可能性が
332 あるため、データソースにおけるいくつかの特徴を考慮する必要がある。その特徴
333 には、データベース間で使用されるコーディングシステムの差異、データ要素の標

334 準化、データが収集される医療環境（例：一次医療、病院、専門医療、リハビリテ
335 ーション）が含まれる。患者、医療提供者又は医療システムにおいては、データ収
336 集又はデータ提供に対する動機づけ（金銭的、社会的、又はその他）が異なる場合
337 がある。また、保険償還の請求手続きも異なる場合もあり、これらは基礎となるデ
338 ータの特性や、調査デザイン及び結果の解釈に影響を及ぼす可能性がある。

339 近年、さまざまな地域で RWD の**連合型データネットワーク (Federated Data**
340 **Network)** が構築されている。複数のデータソースをネットワーク化又はデータをリ
341 ンケージにより利用する場合、調査実施者は施設間又はデータソース間でデータを
342 調和（標準化）させるための手順を検討すべきである（連合型データネットワーク
343 を参照）。いくつかの連合型ネットワークは、科学的な評価及び規制上の意思決定を
344 支援するために特別に設計されており、これらネットワークより提供されるデータ
345 （多くの場合は様々な国々からのデータ）を利用する調査の数が増えている。連合
346 型データネットワークを用いる場合であっても、選択したデータソースの長所と限
347 界を理解することが不可欠である。

348 5.2.1 注目する安全性の検討課題への対処におけるデータソースの適切性

349 調査実施者は、特定のリサーチクエスチョンに対処するために、データソース及
350 びそのデータソースを用いる適切性に対する理解を示すべきである。選択したデー
351 タソースについて、特定のリサーチクエスチョンに対処するためのデータの適合性
352 及びデータの信頼性を含めて理解すること、及び適切な調査デザインと解析の利用
353 を組み合わせることが、間違いのないエビデンスを得る上で重要となる。調査実施
354 計画書を作成する工程においては、調査実施者は、フィージビリティ評価に基づ
355 き、調査に用いることを検討しているデータソースについて、データの適合性及び
356 データソース選択の理論的根拠に関する説明を補足するために、以下の重要な側面
357 について記述すべきである。また、対象とする患者集団に対する調査結果の一般化
358 可能性にそれらが及ぼし得る影響についても記述すべきである：

- 359 • 選択したデータソースが調査の要素をどの程度捕捉しているか（例：変数が捕
360 捉されているか、捕捉されている場合はその完全性の程度）
- 361 • アウトカム又はその他の主要な要素（例：曝露、主要な共変量、組入れ基準／

362 除外基準) をバリデーションすることができるか

- 363 • 選択したデータソースを薬剤疫学調査の目的で使用した経験があるかどうか。
364 これには、提案された調査におけるデータソースの利用をサポートするため
365 に、使用に適合性があることを示すデータの特徴又はその他の要素を実証して
366 いる可能性がある過去の薬剤疫学調査を示した公表文献を参照することも含ま
367 れる。
- 368 • データが利用可能になるまでの時間、データの更新頻度
- 369 • 薬剤の階層化（例：第一選択、第二選択）、フォーミュラリーの判断及び患者
370 への適用範囲等といった、医療システムの構成因子は、疾患の重症度又は他の
371 特性がある医療システムで特定の治療を受けている患者と別の医療システムで
372 同じ治療を受けている患者との間でどの程度異なるかに影響を及ぼす可能性が
373 ある
- 374 • 年齢、社会経済的状態、健康状態、アウトカムのリスク因子、医療システム
375 （民間医療か公的／政府による医療か）等、潜在的な交絡因子として作用する
376 可能性のある重要な患者背景。
- 377 • リサーチクエスチョンに対処するためのデータソースの潜在的限界。

378 5.2.2 主なデータソースの特徴

379 調査の実施にあたっては、利用するデータソースにかかわらず、エビデンスの生
380 成に影響するデータソース固有の情報（例：地理的位置、データが作成された状
381 況、データが収集された期間、データソースに含まれる集団の年齢及び、性別分布
382 等の人口統計学的情報）を入手すべきである。本ガイドラインでは、医薬品の市販
383 後安全性を評価する目的で利用されうる RWD として、例えば、EHR から得られた
384 データ、**診療報酬及び調剤報酬明細書のデータ（レセプトデータ）** (*administrative*
385 *healthcare claims data (claims data)*)、患者レジストリから得られたデータ、患者由
386 来のデータ、健康状態に関する情報を得ることができる他の情報源（インタビュー、
387 郵送調査、コンピュータ又はモバイルアプリケーションによる質問票、**デジタル**
388 **ヘルステクノロジー (DHT, digital health technologies)** を用いた測定等) (デジタル
389 ヘルステクノロジーを参照) から収集したデータ等を想定している。医療環境等

390 の地域的な差異があることから、すべての場合に当てはまるものではないが、医薬
391 品の市販後安全性を評価する目的で利用されうる RWD に含まれる代表的なデータの
392 種類について、一般的な考慮事項を以下に示す。

393 電子診療録 (EHR) データ

394 **電子診療録 (EHR, *electronic health record*)** データは医療機関によって収集され、
395 特定の医療機関内で収集された断片的な治療エピソードを反映している。このた
396 め、その他の医療環境におけるデータは含まれていないことから、患者の完全な病
397 歴を反映していない可能性がある。データの構成要素及びフォーマットが医療機関
398 によって異なる可能性があることから、複数の医療機関のデータを統合する調査に
399 において、データフォーマットの標準化が大きな問題となることが多い。

400 重要な臨床情報は、多くの場合、EHR では非構造化データとして記録されてお
401 り、フリーテキストフィールド (医療関係者のメモ等) 又はその他の標準化されて
402 いない電子的な文書 (PDF ベースの放射線診療録等) として存在する。フリーテキ
403 ストは、EHR ベースのデータソースにおける曝露及びアウトカムのみならず特徴づ
404 け (例: 患者プロフィールのレビュー) に利用される場合がある。データ抽出の効
405 率を高めるため、非構造化データを計算可能な構造化データの形式に変換する既存
406 技術及び新技術 (例: 自然言語処理、画像認識技術、臨床検査結果の評価) を含む
407 様々な手法の活用が増加している傾向にある。

408 複数の医療機関から収集した EHR データを二次利用する場合、利用されるコード
409 (病名、薬剤名、臨床検査項目等) を含め、これらのデータの内容及び形式に差異
410 があれば、調和 (標準化) を図り、その方法を実施計画書に記載すべきである。
411 EHR データは、医療提供者における典型的な医療上のやりとりを収集しているが、
412 実際の医療提供を反映していない場合があり (例: オーダーはされたが調剤又は投
413 与はされていない医薬品)、追加の情報とのリンケージが必要な場合がある (例: 薬
414 局に対する記録)。さらに、特定の種類のプライバシーに関する懸念 (例: 性感染
415 症、薬物使用障害、精神健康状態) を有する患者について、包括的な医薬品使用歴
416 又は健康・医療データを入手することが困難な場合がある。しかしながら、このよ
417 うなデータを収集できなければ、不正確又は不完全なデータになる可能性がある。

418 レセプトデータ

419 レセプトデータベースは、健康保険制度の対象となるすべての個人の医療サービ
420 スが記録されており、大規模なデータである場合が多い。一般的に、健康保険制度
421 内で個人に提供されたすべての医療に対する請求が完全に裁定（すなわち、保険会
422 社又は請求処理業者による最終的な支払い決定）されると、その情報は保険請求の
423 より詳細な情報を含むデータベースに集約される。データベースの中には、オーブ
424 ン（処理中）及びクローズ（支払い済み／却下）の請求情報が混在しているものも
425 あり、調査実施者はこのような場合のデータの動的性質を理解しておくべきであ
426 る。他のデータソースとの紐づけがなければ、医療機関への来院、臨床検査の結
427 果、妊娠に関する調査における出産の転帰、多くの予防接種、事故による負傷及び
428 健康保険でカバーされないその他の治療に関する情報を得ることができないことが
429 多い。このような問題は、健康保険の適用方針や、保険制度外の医療の利用（例：
430 自費・セルフケア治療、労災保険により補償される処置、自動車損害賠償責任保
431 険）等、多くの要因に起因しうる。

432 レジストリ

433 **レジストリ (registry)** は、特定の疾患、状態又は曝露によって定義された集団か
434 ら、事前に収集する項目が規定された一様なデータを収集する体系的なシステムで
435 ある。レジストリは、それらの特徴を表して「患者レジストリ」又は「製品レジス
436 トリ」と表現されることがある。前者では、特定の疾患を持つ患者、妊娠中又は授
437 乳中の人等の特定の集団、先天性欠損症又は分子的若しくはゲノム的特徴等の特定
438 の状態にある個人からのデータ収集に重点を置いている。後者は、調査の依頼者が
439 特定の医療製品又は製品群に曝露された患者のデータを収集するシステムである。

440 すでに確立されたレジストリが、当初の意図とは異なる目的でデータを集めるた
441 めに利用されることがある。調査で既存のレジストリデータを二次的に使用する場
442 合、EHR、レセプトデータ等の二次利用と同様の考慮事項及び目的への適合性の評
443 価を行い、リサーチクエスチョンに答えるためのレジストリの適切性を評価すべき
444 である。例えば、レジストリに登録される集団、収集されたデータ要素（長期性、
445 データ評価の頻度、暦時間等を含む）、データの品質レベル及びガバナンス（データ
446 共有及びデータアクセスに関する側面を含む）を考慮することが挙げられる。ま

447 た、その他の考慮事項として、レジストリの種類及び患者選択の方法が目的とする
448 集団に対する代表性に及ぼす影響（地理的要因、レジストリ登録患者総数、適格患
449 者数、年間新規登録患者数、年間喪失患者数及びその理由等）がある。リサーチク
450 エスチョンに答えるために必要なデータが、既に存在するレジストリにおいて日常
451 的に収集されていない場合は、外部のデータソースとのリンケージ又は他の手段に
452 よる補足的なデータ収集を検討すべきである。

453 場合によっては、新たなレジストリ又は他の調査デザインが必要となる場合もあ
454 る（例：単一群の製品レジストリにおける適切な比較対照集団の必要性、曝露又は
455 共変量の重要な測定項目（治療期間、用量、投与経路等）、無作為化を必要とするよ
456 うな調整不能なチャネリングバイアスへの対処）。

457 **デジタルヘルステクノロジー（DHT）により収集されるデータ**

458 デジタルヘルステクノロジーは、ヘルスケア及びその関連する用途のために、コ
459 ンピューティングプラットフォーム、ネットワーク接続、ソフトウェア及び／又は
460 センサーを使用するシステムである。これらの技術は、一般的なウェルネスへの活
461 用から医療機器としての活用まで、幅広い用途に及ぶ。これらの技術には、一つの
462 医療製品として、医療製品に含まれるものとして又は他の医療製品（機器、医薬
463 品、生物学的製剤）の補助として利用されることを意図した技術が含まれる。ま
464 た、DHT は、医療製品の開発又は調査に使用されることもある。技術の進歩によ
465 り、従来のデータソースを補完するために利用できるデータソースの範囲が拡大
466 し、医療介入の安全性（及び有効性）に関する洞察が得られる可能性がある。これ
467 らの技術は、他のデータソースと同様に目的への適合性評価を受けるべきである。
468 DHT の特定（例：バージョン、ソフトウェア、ハードウェア、製造業者）や、異な
469 る種類のデバイス間でのデータの調和等が必要となる場合がある。データソースの
470 成熟度によっては、より多くの検証作業が必要になる場合がある。

471 **連合型データネットワーク**

472 **連合型データネットワーク（FDNs, federated Data Networks）**は、複数のデータベ
473 ースにまたがるデータ又は結果を統合する分散型の解析が行えるものである。調査
474 に FDN を利用することを選択する場合、FDN による**共通データモデル（CDM,**

475 *common data models*) へのデータの変換及びデータの由来するシステムの差異等、
476 FDN 特有の課題を考慮すべきである。連合型データネットワークの管理（中央集権
477 型又は分散型）についても、調査の運用面に影響を及ぼすため考慮する必要があ
478 る。

479 FDN のフレームワークの下では、複数のデータベースからのデータ又は結果を統
480 合するために、異なる手法を適用することができる。すべての手法に共通する特徴
481 は、データパートナー（FDN を構成する各データベースのデータ保有者）が既存の
482 環境内の電子データに対する物理的・運用的な管理を行っていることであり、デー
483 タ抽出は常に各データパートナーにおいて行われることである。ただし、共通の実
484 施計画書の使用、CDM の使用、データ解析を実施する場所及び方法は FDN 間で差
485 異が存在する。

486 CDM に取り込まれるデータの選択は、利用目的に必要なデータの種類及びデータ
487 の詳細さに応じて最適化される。したがって、CDM 主導型ネットワークにおいて
488 は、個々のデータベースに存在するすべてのデータを含んでいることは稀であり、
489 特定の CDM を採用するネットワークのために選択されたデータ要素は、あらゆる調
490 査の目的及びリサーチクエスチョンに対して十分ではない可能性がある。

491 FDN には、医薬品の安全性の課題への対処に役立つ、以下のような独自の利点が
492 ある：

- 493 • 事前に開発された解析によって又は期待するサンプルサイズを得るために要す
494 る時間を短縮することで調査の対象集団の規模を増大することによって、調査
495 の実施にかかる期間を短縮する。
- 496 • サンプルサイズを大きくすることで、稀な有害事象、稀な疾患及びあまり一般
497 的でない薬剤曝露に関する調査が促進される可能性がある。
- 498 • 複数のデータベースを用いた調査は、異なる集団又は国の間で医薬品の安全性
499 上の問題が存在するか否かに関する追加の知見を提供しうる。また、追加の知
500 見によって薬効に差が生じる原因が明らかにされる可能性があるだけでなく、
501 結果の一般化可能性、情報の一貫性及び推定値へのバイアスが推定に及ぼす影
502 響についての情報を提供する可能性もある。

- 503 • 治療選択肢及び医薬品の利用傾向が医療機関、地域又は国の間で不均一である
504 ことによって、個々の医薬品の効果をより完全に理解することに繋がる可能性
505 がある。
- 506 • 様々な国の専門家が参加し、専門用語、データベースにおけるコード化、調査
507 のプラクティスについて対処、議論することで、薬剤疫学調査の結果の一貫性
508 を高める機会が得られる。

509 データリンケージ（データの紐づけ）

510 データリンケージ（データの紐づけ）は、個々の患者に関するデータの幅及び深
511 さを徐々に拡大するために実施することができる。また、バリデーションの作業を
512 支援するために他のデータソースへのアクセスできるようにするためにも実施され
513 る。がんレジストや死亡レジストリなどのデータソースをレセプトや EHR と紐づけ
514 ることによって、元のデータソースにはないデータを含めることができ、より品質
515 の高い調査につながる可能性がある。データを総合的に理解し、リンケージ及びそ
516 の結果得られるデータについて、それぞれのデータの正確性及びデータの完全性を
517 評価することが重要である。状況によっては、コード化された入力データと、それ
518 に紐づけられた電子フォーマットのカルテレビュー又はテキスト情報が、曝露、ア
519 ウトカム又は共変量の同定に有用である場合がある。

520 概念的には、データリンケージはデータベース内（例：母親と子供のリンケー
521 ジ）又はデータベース間（例：バイタルサインのデータ、バイオバンク）で行われ
522 る場合もある。データリンケージを伴う調査の場合、調査実施計画書には各デー
523 タソース、取得される情報、リンケージの方法並びに経時的なデータリンケージの正
524 確性及び完全性について記述すべきである。追加データの生成（例：インタビュー
525 ー、アンケート調査、コンピュータ化又はモバイルアプリケーションを用いた質問
526 票への回答、デジタルヘルステクノロジーを用いた測定）を伴う調査の場合、調査
527 実施計画書には、データの収集方法、データリンケージの方法、リンケージに使用
528 するデータ要素の説明並びにリンケージが不完全な場合若しくはリンケージされた
529 データソース全体で矛盾するデータが見つかった場合の対処方法について記載すべ
530 きである。

531 5.2.3 データの標準化

532 データの標準化は、連合型データネットワークなどの複数のデータベースを使用
533 する調査に関係する。RWD ソースから得られたデータを標準化する際に考慮すべき
534 課題には以下が含まれる（ただし、これらに限定されない）：

- 535 • 情報源に含まれる情報の種類（例：診断、処置、投薬）
- 536 • RWD ソースの多様性並びにフォーマット及びコーディング言語の一貫性の程
537 度。例えば、異なる標準化及び用語を用いて地域的及び世界的に収集されたソ
538 ースデータの差異
- 539 • 医療システムにおける差異。例えば、ビジネスプロセス及び地域の医療プラク
540 ティスパターン、データベース構造、語彙、コーディングシステム並びにデー
541 タの共有時に患者データを保護するために使用される非識別化手法等。

542 診断、医薬品及び検査データ等のコーディングシステムは定期的に更新されてい
543 る。そのため、コーディングシステム発展／変化に伴うマッピング計画について
544 は、調査計画の策定段階で検討されるべきである。さらに、コードリストには個々
545 の調査の目的、方法及び作成時期が反映されるため、他の調査のコードリストを再
546 利用するには注意が必要である。

547 一部のデータベースにはフリーテキスト／非構造化された構成要素が存在し、こ
548 れは組入れ基準、除外基準、曝露、アウトカム、追跡、及び共変量を定義するため
549 に利用することができる。フリーテキストの各構成要素は、測定対象、測定タイミ
550 ング、治療の設定並びに曝露、アウトカム、共変量等の定義又はアルゴリズムに用
551 いられるコードとしてユーザに提示するため、構造化された表に変換される。非構
552 造化データから構造化変数を作成する工程は実施計画書に記載されるべきである。

553 5.2.4 欠損データ

554 欠損データとは、注目するデータソースに補足されていないデータ値である。デ
555 ータが欠損しうる状況は2つある。1つ目の状況は、収集される想定 of データが収集
556 されなかった場合である。2つ目の状況は、データソースにデータを収集することが
557 想定されておらず、それにより利用できない場合である。EHR システム又は診療報

558 酬及び調剤報酬のレセプトデータベースにおける記録は、患者と医療システムとの
559 やり取りがある場合にのみ生成される。臨床検査結果又は処方等の情報の欠損は、
560 (1) 医療提供者がそれらの指示しなかった場合、(2) それらの指示はされたが実施
561 されなかった場合、(3) 実施されたが、その事実（検査、調剤）が記録されなかつ
562 た場合又は(4) 医療上のやり取りの証拠があり、検査、調剤の等の事実がデータソ
563 ースに保存されたが、データがアクセス可能な形式ではなかった若しくは調査固有
564 の最終的なデータセットが作成された際の変換及びデータのキュレーションの処理
565 中に失われた場合等、異なる状況によって生じる可能性がある。欠損の取り扱い方
566 法の詳細は7項「解析」に記載する。

567 5.2.5 データの品質

568 管理や文書化など、エビデンスを生成する過程の各段階におけるデータの品質に
569 関する基本的な決定要因については、調査実施計画書を最終確定する前に検討する
570 必要がある。データソースによっては、薬剤疫学データは記録、収集及び保管の過
571 程において**品質管理 (QC, quality control)** が厳格ではない場合がある。その結果、
572 不完全なデータ、重要な変数の欠損及び不正確な記録につながる可能性がある。こ
573 のようなデータの品質の欠陥が存在することは、その後の**データのキュレーション**
574 (**data curation**)、データの適用可能性及びデータの追跡性に影響を及ぼす。

575 医薬品の臨床試験の実施の基準 (GCP) と比較して、薬剤疫学データの品質管理
576 及び**品質保証 (QA, quality assurance)** の手順は、薬剤疫学データに特有のガイドン
577 スに従い、満たすべき詳細な品質の基準については実施国又は地域の規制要件に従
578 うべきである (QA 及び QC の詳細については6章「データマネジメント」を参照)。

579 5.2.6 調査実施計画書におけるデータ収集及びデータソースに関する項目

580 調査実施計画書には、使用するデータソース及び対象となるリサーチクエスチョ
581 ンに取り組むにあたって、どのようにデータソースが適合しているかを記載すべき
582 である。また、調査実施計画書には、曝露及びアウトカムの分類に使用するコーデ
583 ィングシステム (例：解剖学的治療化学分類 (ATC)、国際疾病分類 (ICD)、デー
584 タリンケージに使用された方法) を明記すべきである。データ収集の方法と手順も
585 記載すべきである。

586 複数のデータソース又は調査施設からのデータ（例：連合型データ、メタアナリ
587 シス、プールしたデータ）を使用する調査について、調査実施者は、データ間で集
588 団特性、実際の臨床行為及びコーディングがばらついている可能性があることを考
589 慮し、異なるデータソースから試験に用いるデータを入手する方法、及び許容可能
590 な品質でデータを統合する方法について、その論理的根拠及び手順を説明すべきで
591 ある。

592 一次データ収集を伴う調査については、関連する管轄地域の法律及び規制に従っ
593 て、治療中に発生した有害事象の特定、処理及び報告について調査実施計画書に記
594 載すべきである（8章「報告及び提出」を参照）。

595 5.3 対象集団／調査の対象集団

596 対象集団とは、推論を行いたい集団（例：注意欠陥多動性障害を有する12～16歳
597 の小児）のことである。また、調査の対象集団は、リサーチクエスチョンに答える
598 ためにデータが評価される、対象集団を代表するよう意図された集団のことであ
599 る。調査の対象集団は、組入れ基準及び除外基準（例：人口統計学的要因、病状、
600 疾患の状態、重症度、バイオマーカー）によって定義され、特に以下の要素に基づ
601 いて特定される：

- 602 • 調査に組み入れる起点となる日、ルックバック期間（例：新規患者を特定する
603 ため）等の時点
- 604 • 調査の対象集団を選択するために利用された主な変数（4.2章「実施可能性の評
605 価」参照）と、その妥当性を確認する方法（5.5章「バイアスと交絡」参照）
- 606 • 組入れ基準及び除外基準を判定するデータ要素の完全性及び正確性（5.2.2章
607 「データの種類」参照）

608 5.4 曝露、アウトカム、共変量

609 フィージビリティ評価における初期段階において、注目する曝露、アウトカム及
610 び共変量が選択したデータソースにおいて適切に捕捉される可能性が高いことが示
611 された場合、これらの要素を定義し運用の検討を進めるべきである。この工程は、
612 通常、最初のデータベース選択時に開始される **概念的な定義** (*conceptual*

613 **definition**) の作成によって始まる。概念的な定義は、(1)集団選択の条件や、注目す
614 るアウトカム若しくは共変量を定義するための臨床的な基準、又は(2)注目する曝露
615 を定義するための薬物摂取量の測定方法のような、注目する変数に関する最新の医
616 学的及び科学的知見を反映するべきである。また、概念的な定義には、曝露、アウ
617 トカム又は共変量を特徴付けるデータ要素の詳細な説明を含めるべきである。

618 続いて、実施可能性の検討段階で特定された重要なデータ要素を活用し、この概
619 念的な定義を**運用上の定義 (operational definition)** に発展させる。運用上の定義
620 は、データソースから最も完全で正確なデータを抽出するために、概念的な定義に
621 基づいて作成すべきである。EHR 又はレセプトデータを用いる多くの薬剤疫学調査
622 では、構造化データの要素を用いたコードに基づく電子的なアルゴリズムが運用上
623 の定義となる。他の薬剤疫学調査では、非構造化データから関連情報を抽出する、
624 又は構造化データの要素及び非構造化データの要素を組み合わせたアルゴリズムを
625 構築することで、運用上の定義を導き出すこともある。なお、運用上の定義では、
626 必要に応じて、患者調査等の追加のデータ収集を定めることもできる。調査実施者
627 は、曝露、アウトカム及び共変量の定義を作成する際、以下の点を考慮すべきであ
628 る：

- 629 • 曝露、アウトカム及び共変量の概念的な定義を、選択したデータソースにおい
630 て運用できるものに変換することが可能か
- 631 • 運用上の定義が、概念的な定義のすべての要素を適切に捉えているか
- 632 • リサーチクエスチョンに基づき選択されたデータソースにおいて、運用上の定
633 義及びその性能特性（感度、陽性的中度（PPV, positive predictive value）等）は
634 適切であるか（5.6章「データのバリデーション」参照）。

635 概念的な定義は、しばしば**表現型 (phenotype)** と呼ばれる。調査実施計画書に
636 は、運用上の定義（コンピュータ処理可能な表現型と呼ばれることもある）の詳細
637 （コーディングシステム及び理論的根拠を含む）、関連する限界（例：測定バイア
638 ス、代替変数）、誤分類による潜在的な影響、及びこれらの限界を調査デザイン並び
639 に解析によってどのように緩和できるかについての詳細な説明を含めるべきであ
640 る。非構造化データについては、詳細な説明、利用することの理論的根拠、アウト

641 カム／曝露／共変量を特定するための検索基準、コード又は概念のリストを提供す
642 べきである。運用上の定義は、調査実施計画書及び／又は**統計解析計画書 (SAP,**
643 **statistical analysis plan)** に記載すべきである。

644 あるデータソース又は薬剤疫学調査の対象集団に対して作成された運用上の定義
645 は、データベース固有の特性並びに集団間及びデータベース間での疾患疫学のばら
646 つきにより、感度又は特異度の点で他のデータソース又は集団とは異なる場合があ
647 る。他の薬剤疫学調査又は他のデータベースで利用された又はバリデーションされ
648 た定義を活用又は適用する場合は、その適用可能性の根拠を示し正当化されなけれ
649 ばならない。

650 特定の調査のためにデータベースで曝露及びアウトカムを特定する場合、これら
651 の種類の情報に関連するデータは通常コード化されている。データソースを選択す
652 る際には、曝露及びアウトカムを定義するためのコーディングシステムの適切性に
653 ついても確認すべきである。

654 以下の要素については、調査実施計画書の作成において考慮し、詳細に記載する
655 必要がある。

- 656 • データソース／種類及び構造
- 657 • 曝露、アウトカム及び共変量の定義並びにそれらを特定するために利用する方
658 法の作成
- 659 • 運用上の定義の作成とその性能。例えば、時点データ、データの種類（構造
660 化、非構造化）、変数型（カテゴリ値、連続値）、変数型の変換、コードの種
661 類、該当する場合は辞書コードのマッピング（例：ICD-10 から MedDRA）及び
662 性能の評価の仕組み（ゴールドスタンダードの選択）と性能指標等
- 663 • 利用可能なデータ要素及びリサーチクエスチョンに必要なデータ要素のマッピ
664 ング。
- 665 • 過去に利用されたアルゴリズムを注目するデータベース／対象集団に利用する
666 ことの変更の妥当性及び適切性についての文書化
- 667 • 調査の妥当性及び解釈に対する誤分類の潜在的な影響

668 5.4.1 曝露

669 概念的な定義

670 曝露とは、薬剤疫学調査で評価される、注目する医薬品（及び用法・用量）をい
671 う。注目するものは治療であり、無治療、標準治療、別の治療又はこれらの組み合
672 わせに対して比較することがある。

673 曝露の定義は、様々な粒度で作成することができ、例えば、曝露歴の有無、曝露
674 期間、使用者の種類（例：新規使用、既存使用）、リスク期間又はリスクウィンドウ
675 とも呼ばれる曝露ウィンドウ（例：現在の曝露、過去の曝露）、複数の曝露（例：複
676 数の医薬品の使用、予防接種の併用）、治療の切り替え、順序（例：第一選択、第二
677 選択）、用量（例：現在の用量、長期にわたる累積量）等がある。概念的な定義の作
678 成においては、調査デザインの要件及びデータの利用可能性の両方を考慮すべきで
679 ある。曝露の定義には、医薬品の用量、ブランド、剤形、含量、投与経路、時期、
680 頻度及び投与期間（該当する場合）に関する情報も含めるべきである。製品識別情
681 報の一部として製造業者の記述が必要な場合もある（例：異なる製造業者が製造し
682 た同一の有効成分を有する医薬品）。曝露の定義には、当該医薬品又は当該製品クラ
683 スの薬理的又は生物学的特性の理解が必要となる場合がある。

684 曝露の取扱い

685 医薬品の種類、投与経路、治療の設定について

686 概念的な定義を運用上の定義に変換する際には、考慮すべき不確実性があり、そ
687 れらは調査実施計画書の長所と限界に関する議論において正当化されるべきであ
688 る。考慮すべき不確実性の例：

- 689 • 処方された医薬品が必ずしも調剤されるとは限らないこと
- 690 • 調剤された医薬品は、必ずしも使用・投与されるとは限らないこと
- 691 • 患者のコンプライアンス及び投与量を正確に把握する技術
- 692 • 医薬品のサンプル、安価な医薬品、処方箋不要の医薬品、職場で提供される予
693 防接種等、データソースでは補足されない曝露が存在すること
- 694 • 曝露を特定するために使用されるコーディングシステム（例：NDC、RxNorm、

695 HCPCS 等)

696 医薬品の投与環境は慎重に検討すべきである。点滴は、入院時に加えて、民間の
697 診療所又は外来（在宅医療等）でも行われることがあるため、データリンケージが
698 必要となる可能性の観点から、医薬品の投与環境及び治療パターンを考慮すべきで
699 ある。

700 ワクチンについては、製品名、投与スケジュール、他のワクチンとの同時接種、
701 場合によってはバッチ番号又は接種方法及び投与部位に関する詳細な情報が不可欠
702 であり、これらのデータは、ワクチン接種登録情報へのリンケージが必要となる場
703 合がある。

704 医薬品の用量、時期及び曝露の期間について

705 対象となる医薬品への曝露（例えば、用法・用量）は明確に定義し、測定される
706 べきである。医薬品の曝露のタイミング（例：アウトカムの発現時期に対して適切
707 な曝露期間）を考慮すべきである。これは「頓用」又は非処方せん薬が注目する曝
708 露の場合、特に困難である可能性がある。曝露期間を定義する際、曝露の開始日
709 （指標日）を処方日、調剤日又は投与日のいずれとするかを決定する必要がある。
710 患者が処方薬を正確な予定どおりに補充しない場合や、予定より早期に補充する場
711 合があるため、治療におけるギャップ又は医薬品の備蓄があることで、その状況が
712 データに反映されている可能性がある。ある曝露期間が継続しているかを判断する
713 際には、曝露エピソードを構成するための調剤間に許容されるギャップや、複数の
714 曝露エピソード間に許容されるギャップを考慮すべきである。曝露期間を終了する
715 条件も検討し明確に定義すべきである（例：その先 6 か月間に新規処方の記録がな
716 いこと）。その際には、調査に用いられたデータセットには補足されていない可能性
717 がある他の医療環境において、患者に薬剤が処方されている可能性がある等、限界
718 があることに留意する。

719 5.4.2 アウトカム

720 概念的な定義

721 注目するアウトカムの定義は、必要に応じて、疾患の臨床的、生物学的、心理学
722 的及び機能的な概念に基づくべきである。この概念的な定義は、医学的及び科学的

723 な知見を反映するべきである。アウトカムを特定する方法の検討では、真に発現し
724 た事象（以前から有病状態ではなく）として特定できるか、潜伏期間及び増悪を伴
725 う又は再発性エピソードを伴うアウトカムであるか等を考慮すべきである。概念的
726 な定義には、アウトカムを確認するために用いるデータ要素（例：徴候、症状、臨
727 床検査値、放射線検査結果）の詳細な説明を含めるべきである。また、アウトカム
728 の概念的な定義には、診断基準、測定方法及びその品質管理（もしあれば）、測定ツ
729 ール（例：質問票の利用）、計算方法、測定時点、変数の種類、変数の種類の変換
730 （例：量的変数から質的変数）並びにエンドポイントイベントの評価メカニズム
731 （例：エンドポイントイベント委員会の運営方法）についても含めるべきである。

732 アウトカムの取扱い

733 運用上の定義は、調査で利用可能なデータを単独で用いて、調査の目的を達成し
734 うる受け入れ可能な性能で実装できる定義をいう。概念的な定義は、診断・処置コ
735 ード（例：ICD-9-CM、ICD-10、Read、MedDRA）、臨床検査（例：Logical
736 Observation Identifiers Names and Codes [LOINC]）及び臨床検査値、非構造化データ
737 （例：医師の診察記録、放射線検査、病理検査報告）又は質問票のような測定ツ
738 ルを用いて運用される。運用にあたっては、長い年月の間に EHR システムやコード
739 の変化が起こることから、その考慮が不可欠である。非構造化データを用いる場合
740 には、利用した手法及びツール、並びにそれら手法の妥当性の検証に関する説明及
741 び論理的根拠を提供すべきである。

742 単一の診断コードは、除外した診断を示す場合や、特異度が不十分な場合があ
743 る。そこで、傷病名又は診断のみに基づいてアウトカムを運用するのではなく、診
744 断や治療に用いられた医薬品、臨床検査データ及び診療行為の組合せ（例：血栓塞
745 栓症の診断コード+抗凝固剤による治療、アナフィラキシーのコード+エピネフリン
746 の使用）によって、アウトカムの確かな定義が作成しうるかを検討する。アウト
747 カムが複雑な定義となる調査においては、診断を行う医師の専門分野に関する情報
748 が、アウトカムを定義するために利用される情報の質の更なる担保に有用である可
749 能性がある。アウトカムとしての死亡は、患者が死亡時に医療を受けていない限
750 り、健康・医療データに含まれない可能性がある。外部の人口動態統計の資料との
751 リンケージが必要な場合もある。

752 過去に作成された運用上の定義の利用を検討する場合、調査実施者は、疾患又は
753 診断に関する長期的な変化や、最近のデータを用いた検証が必要となりうるコーデ
754 イング手順の変化について検討すべきである。公表されているアウトカムの**症例定**
755 **義 (case definitions)** を利用することでも良いが、特定の RWD データセットで利用
756 可能な情報と必ずしも適合しているとは限らない。過去の調査、試験又は研究で評
757 価された運用上の定義を利用することを提案する場合、理想的には、同じデータソ
758 ース及び類似の対象集団において評価された定義を選択する。さらに、感度、特異
759 度、的中度を確立するために実施された過去の調査、試験又は研究の質は常に評価
760 すべきである。過去の調査、試験又は研究で使用された定義又は他のデータソース
761 でバリデーションされた定義の適用可能性は、調査実施計画書に記載された正当性
762 の説明に基づく外的妥当性の評価に依存する。

763 複数のデータソース（データベース、医療機関、施設）からのデータを用いて調
764 査を実施する場合は、診断のコード化、臨床検査値の基準範囲及び投薬記録等のデ
765 ータソース間のデータの差異を考慮してアウトカムを定義する。また、同じ患者で
766 あっても、レセプトデータの治療開始日と、EHR データの診断日は一致しない場合
767 があるため、これらのデータ要素間において、タイミング及び関係性を完全に理解
768 することが不可欠である。

769 患者体験を測定する指標（例：生活の質（quality of life）、疾患の主観的な重症
770 度）がアウトカムに含まれる場合、調査実施計画書には、アウトカムの指標がどの
771 ように定義され、構築され、バリデーションされたか、及びデータ収集の手順を明
772 記すべきである。データ収集の理由及びデータを生成した医療システムの性質は、
773 利用可能な情報の質及び潜在的なバイアスの存在に影響を及ぼす可能性があること
774 から、調査実施計画書にはこれらも記載すべきである。

775 5.4.3 共変量

776 共変量とは、曝露又は注目するアウトカムのいずれでもない変数であるが、集団
777 を特徴づけるものであるか、調査デザイン又は解析で考慮すべき潜在的な交絡因子
778 又は効果修飾因子であるため、測定される変数である。曝露及びアウトカムと同様
779 に、共変量の定義についても、概念的な定義から、必要に応じて、臨床的、生物学

780 的、心理的及び機能的概念に基づく運用上の定義に発展させる。概念的な定義に
781 は、共変量を構成するために用いるデータ要素の詳細な説明を含めるべきである。

782 潜在的な交絡及び**効果修飾 (effect modification)** は、調査実施計画の作成中に考
783 慮し、計画が立てられるべきである。例えば、人口統計学的変数（例：年齢、性
784 別、人種、民族性）、他の曝露（例：生物学的に活性な薬草）、又は関連する併存疾
785 患による効果修飾の可能性は調査中に記録しておくべきである。また、選択したデ
786 ータソースにおいて、関連する効果修飾因子は利用できるべきである。

787 • **交絡 (confounding)** : 交絡は、関連性の尺度の推定が別の因子の存在によって
788 ゆがめられときに発生する。ある変数が交絡因子であるためには、その変数は
789 因果経路上に存在せずに、曝露及びアウトカムの両方に関連していなければな
790 らない。

791 • **効果修飾** : 効果修飾は、あるアウトカムに対する1つの曝露の効果が、別の変数
792 (すなわち、効果修飾因子) の値に依存する場合に起こる。効果修飾因子は必
793 ずしも因果経路に含まれている必要はない。

794 概念的な定義

795 調査に必要な共変量の定義は明確化されるべきであり、利用予定のデータソース
796 でその共変量が直接運用しうるか判断する。当該データソースで共変量が利用でき
797 ない場合、調査実施者はその共変量の代替が適切か検討することができる。

798 共変量の取扱い

799 概念的な定義から運用上の定義への発展は、曝露やアウトカムの場合と同様に進
800 める。共変量はコホートの特徴づけ、傾向スコアの算出、患者の層別化又はマッチ
801 ング、効果修飾の評価及び交絡の調整に利用される。共変量は通常、注目する曝露
802 の開始より前の期間（ベースライン）において特定及び評価される。ベースライン
803 における複数の共変量は、それぞれ異なる期間に評価を行っても良い。ベースライ
804 ンの共変量を評価するために遡る期間（ルックバック期間）の長さは、コーディン
805 グ又は医療環境の変化、予想される受療行動の頻度、リサーチクエスチョンとの関
806 連性、調査の検出力への影響等の要因を考慮して選択される。共変量は、静的な変
807 数又は時間とともに変化する変数のいずれかとして、観察期間中に評価してもよ

808 い。そのため、各共変量を評価する時期を含めて、共変量の適切な評価は、結果の
809 妥当性にとって不可欠である。これらの共変量に関する情報も利用可能性に応じ
810 て、あるリサーチクエストionsに対する調査に特定のデータベースが適する場合も
811 ある一方で、適しない場合もある。調査実施者は、すべての共変量に対して作成さ
812 れた運用上の定義ついて、コード及び治療の設定を含めて、調査実施計画書に記載
813 すべきである。

814 5.5 バイアスと交絡

815 注目するアウトカムに対する曝露の効果について、根拠が確かで正確な推定を得
816 るためには、調査では 2 つの誤差の原因に対処しなければならない。偶然誤差とは
817 異なり、系統誤差（バイアス、交絡）はサンプルサイズを大きくしても対処できな
818 い。むしろ、調査デザイン、実施及び解析の段階で対処するのが一般的である。疫
819 学的な観点から、バイアス（例：選択バイアス、情報バイアス、デザイン又は測定
820 のエラーに起因するバイアス）及び交絡の概念を区別することは有用である。なぜ
821 なら、これらは異なる機序から生じ、調査デザイン及び解析において異なる方法や
822 手法によって対処しうるためである。デザイン及び解析の段階では、情報バイアス
823 及び組入れ基準／除外基準又は追跡不能に起因する選択バイアス等の潜在的なバイ
824 アスの評価を含めるべきである。また、特に一部のデータ要素を収集又は測定でき
825 ない場合に生じる可能性のある交絡の評価についても含めるべきである。そのた
826 め、欠損データの取り扱いについても、調査実施計画書のデータマネジメントの章
827 （6 章「データマネジメント」参照）又は解析の章（7.1 章「統計解析」参照）にお
828 いて事前に規定すべきである。

829 提案されたデータソースは、バイアス及び交絡を適切に調整しうる重要な因子の
830 情報を取得する上で十分か判断するために評価されるべきである。元のデータソー
831 スでは測定されていない又は不完全に測定されている重要な変数の捕捉を拡大する
832 ために、他のデータソースとのリンケージ及び追加のデータ収集を検討すべきであ
833 る。バイアス及び交絡の原因を検討し、デザインの検討段階において、バイアス及
834 び交絡の影響を評価する計画とともに、それらに対処するための方針の正当性につ
835 いて、調査実施計画書、解析計画書又は最終報告書において示されるべきである。

836 5.5.1 選択バイアス

837 選択バイアスには、紹介バイアス (referral bias)、自己選択バイアス (self-selection
838 bias)、既存使用者バイアス (prevalent user bias)、追跡不能 (loss to follow-up, 時間関
839 連バイアス[time-related bias]) 等の様々な種類がある。これら異なる形態の選択バイ
840 アスは、デザイン段階 (推奨) 又は解析段階のいずれかで対処することができる。

841 選択バイアスとしてよく認められるのが、既存使用者バイアス (prevalent user
842 bias) であり、医薬品の既存使用者が薬剤疫学調査に含まれる場合 (すなわち、調
843 査の追跡が開始される前から一定期間にわたり既に治療を受けていた患者が含まれ
844 る場合) に発生する可能性がある。既存使用者は、その調査では把握されていない
845 初期の薬物治療の「生存者」である。リスクが時間とともに変化する場合、選択バ
846 イアスが生じる可能性がある。例えば、新しい薬剤の利用を開始し、安全性に関す
847 る問題を経験し、その薬剤の利用を中止した患者は、調査に組み入れられないこと
848 があり、その場合、治療によるリスクが過小評価される可能性がある。

849 5.5.2 情報バイアス

850 情報バイアスは、2値変数又はカテゴリー変数の誤分類及び、連続変数の測定誤り
851 がある場合に発生する。曝露がアウトカムに及ぼす効果を正確に推定するためには
852 は、内的妥当性を常に最適化し、重要な変数の誤分類を最小限に抑えるべきである
853 が、調査課題又は規制上の目的によっては、ある程度の誤分類が一部の調査では許
854 容される場合があり、個別の状況に応じて判断する。全体として、変数のバリデー
855 ションの程度 (5.6 章「データのバリデーション」参照) は、必要とされる確実性の
856 程度及び潜在的な誤分類が調査の推論に与える影響によって決定されるべきであ
857 る。3章で議論したとおり、定量的バイアス解析を利用する計画は、バイアスを軽減
858 するための戦略を立てるためにバイアスの方向性及び程度を評価する場合、並びに
859 調査のバイアスが当該調査の解釈にどのように影響を及ぼすか評価する場合に有用
860 である (7章「解析」参照)。

861 5.5.3 不死時間バイアス

862 不死時間バイアス (immortal time bias) とは、注目するアウトカムが生じえないコ
863 ホートの追跡期間をいう。不死時間バイアス及びその他の時間が関係するバイアス

864 の要因を防ぐため、適切な起点とする日の選択が不可欠である。不死時間バイアス
865 及びその他の時間が関係するバイアスは、デザインフレームワークの手法を用いて
866 (4.1 章「リサーチクエスチョン」参照)、適格性及び追跡期間の開始日を含むベー
867 スライン情報の評価を通じて軽減できる可能性がある[4]。

868 5.5.4 交絡

869 通常、調査実施者はリサーチクエスチョンに関連するすべての潜在的な交絡因子
870 を把握することができないことから、未測定 of 交絡又は残差交絡が生じる。薬剤疫
871 学では、一般的に考えられる交絡因子として、人口統計学的要因、適応疾患、疾患
872 の重症度、過去の薬剤使用、併用薬、併存疾患、予後特性、フレイル等があり、調
873 査課題によって異なる。未測定 of 交絡に対処又は評価するため、高次元傾向スコ
874 ア、陰性対照及び調査で用いるデータベースでは測定されていない交絡因子に関す
875 るデータを含む調査（サーベイ）等の外部データソースとのリンケージ等、多くの
876 方法が利用可能である。潜在的な交絡因子の存在及びその影響は、調査デザインの
877 立案段階で検討されるべきである。7章「解析」で議論されるとおり、有向非巡回グ
878 ラフは、変数間の関係を理解し、長期的な追跡を伴う調査における潜在的な交絡及
879 び中間効果の特定並びに定量的バイアス解析を用いてこれらの影響を評価するた
880 めに利用することができる。

881 5.6 データのバリデーション

882 妥当性は、ある概念（変数）が、その調査において運用上の定義によって正確に
883 測定される程度の指標である。曝露、アウトカム及び主要な共変量のバリデーショ
884 ンは、薬剤疫学調査の内的妥当性を確保するうえで重要である。バリデーションに
885 は様々な手法があり、それはデータソース固有のものである場合がある。それらの
886 手法には、完全な検証、部分的な検証、臨床専門家によるレビュー、レセプトデー
887 タのレビュー又は患者プロフィール履歴のレビューが含まれる。バリデーション
888 は、要求されるエビデンスレベルに見合ったものであるべきであり、例えば、すべ
889 ての潜在的なケース又は非ケースのアウトカム変数を検証する場合又は運用上の定
890 義によるケース及び非ケースの特定の性能を検証する場合等がある。薬剤疫学調査
891 において頻繁に使用されるデータベースでは、主要な変数について文書化されたバ

892 リレーションが以前に実施されている可能性がある。しかしながら、以前のバリデー
893 ーションの結果を外挿する場合には、今回計画している調査の対象集団と外挿を検
894 討しているバリデーショナルの対象集団における有病率、組入れ基準及び除外基準に
895 おける差異、リスク因子の分布及び解析並びにその後の医療、処置及びコード化に
896 対する変更の影響を考慮すべきである。調査の依頼者は、部分的な検証か完全な検
897 証か又は以前に検証された定義の採用等、提案するバリデーショナルの手法について
898 規制当局と協議し合意するため、早期にやり取りを行うべきである。バリデーショ
899 ンの手法の妥当性は、データソース、対象集団、バリデーショナルの実施期間、性
900 能、参照する基準及び要求されるエビデンスレベルを考慮して提案された運用上の
901 定義の適用可能性に関する考察を含めて説明すべきである。

902 バリデーショナルは、独立した実施計画書に基づいて実施されるべきである。運用
903 上の定義をバリデーショナルする際には、報告すべき指標（例：感度、特異度、
904 PPV、陰性的中度（NPV, negative predictive value）、カッパ統計量）を事前に規定し、
905 それらの測定方法を記述する。運用上の定義を選択し、内的妥当性を担保するため
906 の適切なバリデーショナルのアプローチを特定する際には、対象者の偽陽性と偽陰性
907 の間のトレードオフを考慮すべきである。例えば、対象者がまれな場合、感度の高
908 い運用上の定義を選択した上で、すべての潜在的な症例をバリデーショナルする必要
909 があるかもしれない。

910 複数の運用上の定義を検討している場合は、調査デザインの検討段階でバイアス
911 解析を利用してそれぞれの性能を評価及び定量化すべきである。これは、調査の解
912 析段階で行われる一般的な感度解析とは異なるものである。

913 6 データマネジメント

914 薬剤疫学調査におけるデータ品質の管理は、データソースや調査結果の利用予定
915 を含む様々な要因に依存する。データマネジメント及び／又はデータのキュレーシ
916 ョンの計画は、調査を開始する前に策定するべきである。品質保証及び品質管理
917 （QA/QC）の計画は、解析が実施される前に策定されるべきである。また、品質
918 に影響する因子（例：データ保有者によるデータマネジメント、データの品質欠
919 陥、不適切なデータ処理及び解析又は不十分なトレーニング）を特定し、調査の完

920 全性 (integrity) を保つために対処すべきである。満たすべき詳細な品質基準は、実
921 施国又は地域の規制要件に従うべきである。

922 データセットの提出が規制上の要件である場合、規制当局による審査を容易にす
923 るために、状況、内容、ファイル構成及びファイルの作成手順に関する説明を含め
924 るべきである。データセットは、提出される地域の関連する規制上の要件に従って
925 保管されなければならない。データ及び文書を新しい媒体又は新しいフォーマット
926 に移行する場合は、長期的な可読性を確保し、完全性を維持するために検証すべき
927 である。

928 データマネジメント計画

929 データの品質保証プロセス、方針及び手順は、データ品質に対する潜在的な懸念
930 に対処すべきである。潜在的な懸念には、解釈又はコーディングの誤り、データ入
931 力や転送又は変換の精度による誤り、意図の誤り、不十分なトレーニング、データ
932 の完全性、**データの一貫性 (data consistency)** 等が含まれる。

933 データの保存、管理及び統計解析ソフトウェアに関する説明は調査実施計画書に
934 記載すべきである。解析データセットの取得、検証及びデータの完全性を推進する
935 ために利用されたすべての手順は、再現 (replicate) できるように十分詳細に記録す
936 べきである。権限を付与された個人にアクセスを制限することで、データのセキュ
937 リティを常に維持すべきである。

938 品質保証と品質管理

939 6.1 データ保有者

940 データ保有者が実施する QA 及び QC の手順には、データ収集及び管理の信頼性の
941 確保に関する次の情報が含まれうる：該当するデータ収集期間中にデータ保有者に
942 よって実施されたデータのエラー修正又はデータ判定方針の変更の頻度及び種類、
943 データの品質及び／又は妥当性を検討する査読付きの文献、注目するアウトカムに
944 関連する調査期間を通じたコーディング手順の更新及び変更（例：国際疾病分類
945 ICD コード）、調査期間中の主要なデータ要素の変更及びその変更が調査に及ぼす潜
946 在的な影響、長期的な欠損の程度（言い換えれば、注目する特定の変数に関して入
947 手できないデータの割合）、これらの問題に対処するために用いた手順（例：除外、

948 代入) 等。

949 6.2 調査実施者

950 データ保有者がデータの管理を維持し、基礎となるデータの品質について責任を
951 有している一方で、調査実施者は、透明性を確保し、データの長所/限界について
952 理解し、規制当局が要求する品質の基準を満たすために、QC/QA の手順をデータ
953 保有者と調整する責任を有する。さらに、調査実施者は、すべてのデータクリーニ
954 ング、データ処理、解析データセットに関する管理及び品質保証に対する責任を有
955 する。十分な品質保証の必要性と、特定の目的に対する合理的な財源支出の balan
956 スをとるために、リスクに基づく品質保証へのアプローチが推奨される。データの
957 信頼性及びデータの適合性を判断するために不可欠な問題は、調査実施計画書で扱
958 われるべきであり、データ収集、データのキュレーション及び調査特有の最終的な
959 データセットへの変換のための QA/QC の手順が含まれる。

960 調査実施者は、文書化された QA/QC システムの実装及び維持に対する責任を有
961 する。これは、調査実施計画書、各地域の法律、倫理的配慮及び適用される規制上
962 の要件を遵守して、調査が実施され、結果が生成され、報告されることを保証する
963 ためである。これらの工程の文書化には、データソースから最終的な調査の解析用
964 データセットに至るまでのデータの追加、削除又は改変に関する電子的な文書化
965 (すなわち、メタデータに基づく監査証跡、品質管理手順) が含まれるが、これら
966 に限定されるものではない。調査実施者は、特定の調査を実施するためのデータの
967 変更及びその変更の潜在的な影響を文書化すべきである。解析プログラムの作成の
968 品質保証及び品質管理の方法は、調査実施計画書に記載すべきであり、例えば、プ
969 ログラムコードの検査及び/又はプログラムコードの複製の工程又は以前に QA/
970 QC が実施されたプログラムコードが使用されているか否か等を記載すべきである。

971 7 解析

972 解析戦略には、調査の目的に対処するための記述的解析及び推論的解析を含み、
973 バイアス及び交絡の潜在的な要因に対処する。さらに、未測定又は誤測定による交
974 絡及びその他のバイアスの要因についての実証的な評価も含めるべきである。SAP は
975 事前に規定され、フィージビリティ評価から得られた情報を反映し、調査目的を満

976 たすように作成されるべきである。SAP の概要は、調査実施計画書に記載されるべきである。完全な SAP は、独立した文書として又は調査実施計画書における詳細な章として提供する。調査実施計画書と SAP は密接に関連していることに留意しつつ、選択した手法について規制当局と協議することが推奨される。SAP には、結果の確からしさを担保できるよう、調査の再現を可能にするための十分な詳細を提供すべきである。

982 調査によっては、データ主導の解析が行われることがあるが、事前に規定された解析と事後的な解析を区別することが重要である。共変量選択のために用いた解析等の事前に規定された解析は、調査実施計画書及び解析計画において文書化されるべきであり、計画からの逸脱は最終報告書に文書化されるべきである。事後的な解析は、調査の結果を解釈する一助となるようデータを観察して実施されることが多いが、事後的な解析についても、最終報告書に記載し、その実施については正当性が示されるべきである。また、調査実施者は、調査の実施中に行われる解析（例：患者の集積、記述的解析、推論的解析、感度解析、定量的バイアス解析）のスケジュールを作成することを検討する。

991 複数のデータベースを用いた調査を実施する場合又は FDN を使用する場合は、解析手法が検討中のデータベース又は FDN の種類に影響されるため、事前の計画が必要である。メタアナリシスの手法が求められるような、各データソース又は FDN で実施される独立した関連性の解析など、解析計画ではこのような特殊な課題についても考慮する必要がある場合がある。

996 7.1 統計解析

997 7.1.1 主要解析

998 解析は、注目する疫学的指標（例：リスク差又は率差、リスク比又は率比）の偏りのない推定を目指すべきである。解析に関する章では、仮定及び条件を含め、統計解析のために選択した手法の説明及び正当性を記載すべきである。

1001 以下の側面及び要素を含めることを検討する：記述的解析、サブグループ解析、解析に必要な推定の方法及びそれに関連する仮定、予想される調査の規模／検出力／統計学的な精度の推定、交絡及びバイアス（例：誤分類、選択バイアス、情報バ

1004 イアス、時間が関係するバイアス及び結果の妥当性への影響)を調整する計画、調
1005 査の対象集団の比較可能性の評価、感度解析、第一種の過誤の制御(例:逐次解
1006 析)、代表性の評価、欠損データの取扱い方法等。

1007 機械学習又はその他の導出手法の使用を検討する場合、使用されるコンピューター
1008 アルゴリズムの仮定及びパラメータ、アルゴリズムを構築するために利用された
1009 情報のデータソース、アルゴリズムが教師あり(すなわち、専門家による入力及び
1010 レビューの利用)又は教師なしであるかどうか及び手法の検証に関連する指標を明
1011 記する。

1012 7.1.2 欠損データ

1013 調査実施者は、基礎となるデータにおける情報の有無の理由を理解した上で、調
1014 査実施計画書及びSAPを作成すべきであり、欠損データに対処するためのデータリ
1015 ンケージ及び/又は代入を検討し、欠損データの程度が調査結果に及ぼす影響に対
1016 処すべきである(5.5章「バイアスと交絡」を参照)。欠損データを特徴付けるため
1017 に記述的解析を含めるべきである。調査結果及び重要な共変量に関する統計解析の
1018 基礎となる欠損データに関する仮定(例:ランダムな欠損(missing at random)、ラ
1019 ンダムではない欠損(missing not at random))は担保されるべきであり、欠損データ
1020 の意味を検討する。解析では、SAPに記載された方法に従い欠損データに対処すべ
1021 きである。欠損データの程度及び調査結果に及ぼす影響について説明されるべきで
1022 ある。

1023 7.1.3 感度解析

1024 感度解析を計画する際には、各解析の長所及び限界とともに、各解析の理論的根
1025 拠を示すべきである。感度解析では、デザイン、estimands、曝露の定義、アウトカ
1026 ム定義、欠損データ、選択したデータソースの限界等の解析に関連する重要な仮定
1027 を変化させた場合の効果が検討されるべきである。この解析により、認められた不
1028 確実性の程度を考慮した上で、調査結果をより適切に解釈することが可能となる。

1029 定量的バイアス解析では、潜在的なバイアスが関連性の尺度/指標に及ぼす影響
1030 を評価する。調査実施計画書には、バイアスを定量化するために使用される指標
1031 (例:感度、特異度、PPV、NPV)を事前に規定し、注目する変数の妥当性を検証

1032 する際に、その指標がどのように測定されるかを記述すべきである。バイアスを調
1033 整した効果推定値の精度は、信頼区間を用いて定量化されるべきである。これらの
1034 解析は、調査結果の解釈を容易にする可能性がある。

1035 8. 報告及び提出

1036 8.1 有害事象、副作用及び製品品質に関する苦情の報告

1037 調査実施中に確認された有害事象、副作用及び製品品質に関する苦情は、関連す
1038 る規制当局への報告が必要となる場合がある。報告要件は、規制当局の報告要件の
1039 違いにより、当事者（例：製造販売業者（MAH）、他の調査の依頼者又は申請者、
1040 調査実施者、独立した研究グループ）及び地域によって異なる場合がある。ICH
1041 E2D ガイドライン（承認後の安全性情報の取扱い：緊急報告のための用語の定義と
1042 報告の基準について）は、有害事象及び副作用の個別症例安全性報告に関する MAH
1043 のための考え方を示している。その他の報告要件（及び ICH E2D の適用範囲外の関
1044 係者）については、適用される法令及び必要に応じて適切な医薬品安全性監視ガイ
1045 ドラインを参照する。

1046 8.2 規制当局に提出する調査の文書書式と内容

1047 調査の依頼者は、要求される調査資料及び提出のスケジュールについて規制当局
1048 と協議すべきである。これらの文書は規制当局によって異なる場合があり、フィー
1049 ジビリティ評価、調査実施計画書、SAP、中間及び最終報告書等を含むことがあ
1050 る。規制当局が要求する所定の書式及び内容の指定がない場合、調査の依頼者は、
1051 ISPE/ISPOR の HARmonized Protocol Template to Enhance Reproducibility（治療効果に
1052 関する仮説を評価するリアルワールドエビデンス研究の再現性向上に向けた調査プ
1053 ロトコルテンプレート：HARPER）[1, 5]等、アカデミアによって開発されたフレー
1054 ムワークを活用することができる。

1055 9 調査の資料並びに結果の公表及び伝達

1056 透明性の観点から、科学的な意見交換を支援し、再現可能な調査を実施できるよ
1057 うにするために、規制要件で義務付けられていない場合でも、調査実施者は調査の

1058 資料を公開することが推奨される。調査実施計画書及び SAP は、調査開始前に適切
1059 なレジストリで公表され、また結果報告書についても、調査終了後に公表されるこ
1060 とが望ましい。調査結果の公表及び伝達のためのさらなる手段としては、科学的フ
1061 ォーラムへの規制目的以外での提出、科学的出版物としての公表、患者又は医療関
1062 係者向けのコミュニケーション等が挙げられる。

1063 医学文献での報告を推奨するガイドラインがいくつか存在する。それらには、医
1064 学雑誌編集者国際委員会 (ICMJE) が作成した「The Reporting of studies Conducted
1065 using Observational Routinely collected health Data (RECORD) Statement」、
1066 「RECORD-PE」及び「Recommendations for the Conduct, Reporting, Editing, and Publication of
1067 Scholarly work in Medical Journals」が含まれる。また、調査の内容を公表する際に
1068 は、その公表内容が結果報告書の内容と整合するように要約すべきである。公表バ
1069 イアスを避けるために、調査の目的又は仮説に関して、否定的又は結論の出ない結
1070 果が得られた場合であっても公表することが望ましい。

1071 調査の結果は、調査参加者（例：一次データ収集を用いる場合）、一般市民及び患
1072 者に伝えるべきであり、調査結果及びその意味合いを認識し理解できるようにす
1073 る。公表は、関連する安全性情報及び調査の限界を含めて、客観的かつバランスの
1074 取れた販売促進を意図しない方法で、調査全体の結果の事実に基づく要約を含める
1075 べきである。

1076 10 調査の文書及び記録の保管

1077 調査の計画、実施及び結果に関する主要な文書及び記録は、適用される基準及び
1078 各国の要件に従って保管すべきである。市販後の医薬品安全性評価に係る調査で
1079 RWD を利用する際の主要な原則は、GCP（特に一次データ）及び Good
1080 Pharmacoepidemiological Practice（特にデータの二次利用）と同様とする。

- 1081 • すべての調査の情報、文書及び記録は、正確な報告、解釈及び検証が可能な方
1082 法で記録、取り扱い、保存及び保管されていること。また、適用される個人情
1083 報保護法に従って機密性及び患者のプライバシーを確保しなければならない。
- 1084 • 調査の文書の完全性を確保し、偶発的又は早期の喪失を防ぎ、不正アクセス、
1085 改ざん、廃棄、開示又は流布を防止し、監査証跡を確実に維持するためのシス

- 1086 テムが整備されていること。
- 1087 • 調査の開発、実施及び報告に関する文書のあらゆる側面の品質を保証する手順
1088 を備え、必要とされるシステムが構築されていること。
- 1089 • 調査情報は、主要な記録システムが最新の状態に維持されていること及びすべ
1090 ての主要文書が適切に保管されていることを保証するために、リスクに基づく
1091 品質チェック（risk-based quality checks）又はレビュープロセスを用いて、規制
1092 当局からの要請（例：社内又は規制当局の査察）に応じて容易に入手又は直接
1093 アクセスできる状態にすること。
- 1094 • すべての情報は、少なくとも適用される規制要件で要求されている期間にわた
1095 り保管すること。

1096 11 特定の集団における検討事項

1097 特定の（特別な）集団は、製造販売承認前の臨床試験に登録されないことが多
1098 く、妊婦、授乳婦、乳児、小児、青年／若年成人、高齢者、免疫不全患者、障害者
1099 及び／又は稀な疾患を有する者等が含まれる。市販後の薬剤疫学調査では、これら
1100 の集団における医薬品のベネフィット／リスク評価の裏付けとなる重要な情報が得
1101 られる可能性がある。このような集団を対象とする調査では、調査の対象集団を定
1102 義する際に、あらゆる調査に適用される他の検討事項（曝露、交絡因子、アウトカ
1103 ムの定義等）に加えて、特有の課題に関する検討が必要となる場合がある。課題の
1104 例として、希少疾患におけるサンプルサイズの小ささ、高齢者における複数の併存
1105 疾患及び多剤併用、免疫不全患者における症例及び疾患の特徴（例：罹病期間、重
1106 症度）を特定することの難しさ等が挙げられる。

1107 11.1 妊娠に関する調査

1108 妊娠に関する調査におけるデータの二次利用の具体的な課題には、妊娠の特定、
1109 転帰の複雑さ、妊娠年齢又は受胎日、母体や出生児の転帰を特定するための検証さ
1110 れたアルゴリズムの必要性等がある。このような課題を解決するためには、データ
1111 ソース内でのリンケージ（例：母親と子供のリンケージ）、出生レジストリ等の補完
1112 的なデータソースが必要となる場合がある。妊娠レジストリでは、曝露時期、妊娠

1113 年齢、転帰及び共変量について、より詳細な臨床情報を得ることができるが、参加
1114 者の登録及び維持が困難であること、選択バイアスがあること等の課題がある。

1115 「曝露あり」と「曝露なし」の二元的アプローチでは、妊娠中の曝露のパターン
1116 が反映されないため、妊娠 3 カ月ごとに異なるリスクに対処するアプローチを検討
1117 するべきである。リスクウィンドウの定義、受胎及び妊娠開始日の測定、医薬品の
1118 利用パターン（開始日、終了日、用量、頻度、期間等）に注意を払うべきである。
1119 受胎日から推定され得る妊娠年齢の妥当な推定は、曝露の時期を決定するために重
1120 要であり、超音波検査、検査データ等と紐づけたデータの利用可能性が求められる
1121 ことがある。妊娠中の曝露時期に関する正確な情報は、重要な曝露期間の特定及び
1122 特定の影響の生物学的妥当性を知るのに役立つことがある。特に半減期が長い製剤
1123 では、妊娠直前の曝露情報も重要となる場合が多い。

1124 調査におけるアウトカムには、母体の健康に影響を及ぼす妊娠中の転帰（妊娠高
1125 血圧腎症、妊娠糖尿病等）、自然流産、出産／新生児の転帰、出生後数年にわたる子
1126 どもの発達に関する転帰が含まれる。調査実施計画書には、重症度（例：重大な先
1127 天性異常）を含め、注目するアウトカムを定義するための事前の基準を明記し、ま
1128 た、妊娠中の有害な転帰の多くは妊娠期間中に大きく変動することに考慮する。妊
1129 娠に関する調査のアウトカム特定には特有の課題があるものの、標準的な分類シス
1130 テムの使用を検討すべきである。早産及び"small for gestational age"（妊娠期間に比し
1131 て身体が小さいこと）は、レジストリにおいては確実に利用可能であるが、診療報
1132 酬及び調剤報酬明細書のデータでは診断コードにより特定されるか又は妊娠期間及
1133 び出生時体重のデータを用いて計算される場合がある。データ収集方法にもよる
1134 が、先天性欠損症のレジストリは、生児出生、死産／胎児死亡、選択的人工妊娠中
1135 絶について既に判定されているため有用である。また、重大な先天性奇形について
1136 は、母親の記録、乳児の記録又はその両方に記録されることがある。

1137 妊娠に関する調査におけるバイアス及び交絡には、家族歴及び適応による交絡が
1138 含まれるが、これらに限定されるものではない。解析計画では、妊娠の時点に関連
1139 した時間とともに変化する共変量を考慮すべきである。

1140 12 用語集

1141 この用語集は、ICH ガイドラインに基づいて作成された定義に基づいており、規
1142 制関連文書によって補足され、その後、専門学会のベストプラクティス文書や文献
1143 など、他の情報源から引用された規制外のベストプラクティス文書に基づいて作成
1144 されている。

診療報酬及び調剤報酬明細書のデータ（レセプトデータ） Administrative healthcare claims data (claims data) :

患者が医療システムを利用し、その治療についての医療提供者への支払いに対する払い戻しから生じるデータ。

バイアス Bias:

結果における真の値からの体系的な偏り

症例定義 Case Definition:

疾患の医学的及び科学的知見を反映している、疾患の臨床的、生物学的、心理学的、機能的な状態の概念。

共通データモデル Common Data Model:

いくつかのデータベース又はデータセットを横断的に組み合わせた分析を可能にするために、特定の調査、試験又は研究とは独立に、生データを共通の構造、フォーマット及び用語を標準化するための仕組み。構造及び内容を標準化することで、データ全体で標準化されたアプリケーション、ツール及び方法を使用して、幅広い課題に答えることを可能にする。

概念的な定義 Conceptual Definition:

一般的又は定性的な用語で、調査の構成概念（例：曝露、アウトカム、共変量）又は特徴を説明すること。

交絡 Confounding:

交絡は、曝露とアウトカムの両方に関連する因子、かつ曝露とアウトカムの因果経路にない因子の存在の結果として生じる。交絡は、曝露とアウトカムの関連において観察される曝露効果を歪める。

データの正確性 Data Accuracy:

規定の条件下（又は、特定の方法で測定した場合）で、標準値又は既知の真値に対する測定値の近さの程度。

データの完全性 Data Completeness:

必要なデータが存在していること。

データの一貫性 Data Consistency:

医療施設、部門、施設内のユニット、医療提供者又はその他の評価者間のデータが同

じ傾向にあること。

データのキュレーション Data Curation:

特定のリサーチクエスションの統計解析を目的とした元データの整理。

データのキュレーションには、データ抽出(複数のデータソースを含む)、データセキュリティ処理(匿名化又は非匿名化、データの破損、漏洩、盗難、改ざん又は不正アクセスからの保護)、データクリーニング(編集の確認及び外れ値処理、データの完全な処理)、データ変換(共通データモデル、標準化、自然言語処理、医療コード化、生成変数の計算)、データ品質管理、データ転送及び保管が含まれるが、これらに限定されるものではない。

データ保有者 Data Holder:

公共部門の機関及び国際機関を含む法人又はクエスションにおける特定のデータに関するデータの主体ではない自然人であって、適用法令に従って、特定の個人データ又は非個人データへのアクセスする権利又は共有する権利を有するもの。

データの真正性 Data Provenance:

データ(データベース、文書又はリポジトリ)の出所を説明し、そのデータが現在の場所にどのようにして、なぜ到達したのかを説明する監査証跡。

データの適合性 Data Relevance:

主要なデータの要素(曝露、アウトカム及び共変量)が利用可能であること及び調査に代表性のある患者が十分に存在していること。

データの信頼性 Data Reliability:

データの正確性、データの完全性、データの真正性及びデータの追跡性が含まれる。

データの追跡性 Data Traceability:

解析結果(試験報告書の表、一覧表、図)、解析データセット、集計データセット、元データとの関係を把握することができること。

デジタルヘルステクノロジー Digital Health Technology (DHT):

ヘルスケア及びその関連する用途のために、コンピューティングプラットフォーム、ネットワーク接続、ソフトウェア及び/又はセンサーを使用するシステムである。これらの技術は、一般的なウェルネスへの活用から医療機器としての活用まで、幅広い用途に及ぶ。これらの技術には、一つの医療製品として、医療製品に含まれるものとして又は他の医療製品(機器、医薬品、生物学的製剤)の補助として利用されることを意図した技術が含まれる。

効果修飾 Effect Modification:

効果修飾は、あるアウトカムに対する1つの曝露の効果が、別の変数(すなわち、効果修飾因子)の値に依存する場合に起こる。効果修飾因子は必ずしも因果経路に含まれている必要はない。相互作用は、ある結果に対する2つの曝露のアウトカムへの因果効果に関心があり、どちらかの曝露の効果が、もう一方の曝露の値にどのように依存するかに関心がある場合に起こる。

電子診療録データ Electronic Health Record Data:

EHR システム内に含まれる個々の患者記録の集合体。典型的な EHR には、患者の病歴、診断、治療計画、予防接種日、アレルギー、放射線画像、薬局記録、臨床検査及び検査結果が含まれる。

Estimand:

調査の目的によって提起された臨床上の疑問を反映した治療効果の正確な説明。これは、集団レベルで、比較されている異なる治療条件下の同一患者におけるアウトカムは何かを要約するものである。

曝露 Exposure:

提案された調査で評価されている医薬品又は対象とする用法用量である。

連合型データネットワーク Federated Data Network:

一連の分散化され、相互接続されたノード。これは、データがそのノードを離れることなく、ネットワーク内の他のノードによってデータを照会又は分析することを可能にするものである。FDN の例としては、DARWIN EU、Sentinel、CNODES、OHDSI、MID-NET などがある。

医療レセプトデータ Medical Claims Data:

治療等の医療費の請求のために保険会社に提出されたレセプトに関する情報をまとめたもの。レセプトデータでは、診断及び治療の識別のために、世界保健機関の国際疾病分類コード(ICD-CM)診断コード等の標準化された医療コードを使用されている。

医薬品 Medicine:

疾患の診断、治癒、緩和、治療又は予防に使用することを意図した物質又は物質の組合せ。

運用上の定義 Operational Definition:

調査実施者が特定の調査における構成概念を測定するために従う、データ固有の操作又は手順。

患者体験データ Patient Experience Data:

個人によって収集され、疾患又は状態についての患者の経験に関する情報を提供することを意図したデータ。患者体験データは、以下に関する患者の経験、視点、要望、及び優先事項(ただし、これらに限定されない)を把握する情報として解釈することができる。1) 患者の症状及び疾患の自然史、2) 患者の機能及び生活の質に対する症状の影響、3) 患者の治療経験、4) 患者にとって重要な転帰となるインプット、5) 転帰及び治療に対する患者の好み、6) 患者によって定義される問題の相対的な重要性。

表現型 Phenotype / Phenotype Algorithm:

疾患 (例: 2 型糖尿病)、血圧測定値、血糖値、抗生物質の処方等、健康又は医療に関連する観察可能で測定可能な情報。曝露から転帰に至るまで、あらゆる患者の特徴を定義するために使用できる。症例定義を、調査で実行可能なアルゴリズム (EHR から臨床データの要素へのクエリーを含む) に変換するのが、フェノタイプピング・アルゴリズムである。これらのアルゴリズムは、臨床コード (例: ICD-10、SNOMED) を用

いて医療記録からデータを特定し抽出する。これらは「電子的表現型」または「計算可能な表現型」とも呼ばれる。

確からしさ Plausibility:

信憑性、真実性、もっともらしさ。

一次データ収集 Primary Data Collection:

現在実施している調査のために収集されたデータ。

品質保証 Quality Assurance (QA):

調査の実施、データ作成、文書化（記録化）及び報告が、調査実施計画書及び本基準を遵守していることを保証するために設定された計画的かつ体系的な活動。

品質管理 Quality Control (QC):

調査関連の活動の質に求められる事項を充足しているか否かを検証するために調査の品質保証システムの一環として行われる実務的な手法及び活動。

定量的バイアス解析 Quantitative Bias Analysis:

関連性の尺度に影響する系統的誤差に関連する方向、大きさ、及び不確実性を定量的に推定する方法に適用される包括的な用語である。

リアルワールドデータ Real-World Data (RWD):

様々な情報源から日常的に収集される患者の健康状態及び／又は医療提供に関するデータ。

RWDの例としては、電子健康記録（EHR）から得られたデータ、医療請求及び請求データ、製品及び疾患登録から得られたデータ、患者が作成したデータ（モバイル機器、ウェアラブル等）、及び健康状態について情報を提供できる他の情報源から収集したデータ（例：特定の保健システムで収集された遺伝子およびその他の生体分子表現型データ）が挙げられる。

リアルワールドエビデンス Real-World Evidence:

RWDの分析から得られた医薬品の使用法および潜在的な利益またはリスクに関する臨床的証拠。

レジストリ Registry:

特定の疾患、状態又は曝露によって定義された集団から、あらかじめ規定されたようなデータを収集する組織化されたシステム。

データの二次利用 Secondary Use of Data:

既存のデータを収集目的とは異なる目的のために使用すること。

標準治療 Standard of Care:

特定の病気又は状態に対する適切な治療法として、医療専門家に受け入れられ、医療従事者に広く用いられている治療法のことであり、ベストプラクティス、標準的な医療、標準的な治療とも呼ばれる。

統計解析計画書 Statistical Analysis Plan:

統計解析計画書とは、調査実施計画書に記されている解析の主要な特徴のより技術的な詳細を述べた文書であり、主要変数、副次変数、その他のデータに関する統計解析を実行するための詳細な手順を含むものである。

標的試験 Target Trial:

実施可能な場合には、注目する課題に答えうる、仮想的なランダム化試験。

1145

NOTE

1146 **13 関連する規制ガイドライン**

- 1147 European Medicines Agency. Clinical Trials Regulation (EU) NO 536/2014; 2014 Apr.
- 1148 European Medicines Agency. Guideline on Good Pharmacovigilance Practices (GVP) –
1149 Module III. Revision 1; 2014 Sep.
- 1150 European Medicines Agency. Guideline on Good Pharmacovigilance Practices (GVP) –
1151 Module IX. Revision 1; 2017 Oct.
- 1152 European Medicines Agency. Guideline on Registry Based Studies. Amsterdam (NL); 2021
1153 Oct.
- 1154 European Medicines Agency. Guideline on the Content, Management and Archiving of the
1155 Clinical Trial Master File (Paper and/or Electronic). London (GB); 2018 Dec.
- 1156 European Medicines Agency. ICH E2D Post-Approval Safety Data Management. London
1157 (GB); 2004 May.
- 1158 European Medicines Agency. ICH E8 (R1) General Considerations for Clinical Studies.
1159 Revision 1. Amsterdam (NL); 2021 Oct.
- 1160 European Medicines Agency. ICH E9 (R1) Addendum on Estimands and Sensitivity Analysis
1161 in Clinical Trials to the guideline on statistical principles for clinical trials. Amsterdam (NL);
1162 2020 Feb.
- 1163 European Medicines Agency. ICH Guideline for Good Clinical Practice E6(R2). London
1164 (GB); 2016 Dec.
- 1165 European Medicines Agency. ICH Topic E9 Statistical Principles for Clinical Trials. London
1166 (GB); 1998 Sep.
- 1167 European Medicines Agency. Patient experience data in EU medicines development and
1168 regulatory decision-making. Outcome of the workshop on 21st September 2022. 17 October
1169 2022, EMA/786952/2022.
- 1170 European Network of Centers for Pharmacoepidemiology and Pharmacovigilance (ENCePP).
1171 ENCePP Checklist for Study Protocols. Revision 4. EMA; 2018 Oct.
- 1172 European Network of Centers for Pharmacoepidemiology and Pharmacovigilance (ENCePP).
1173 Guide on Methodological Standards in Pharmacoepidemiology. Revision 11. Amsterdam
1174 (NL): EMA; 2023 July.
- 1175 FDA, United States. Draft Guidance for Industry: Real-World Data: Assessing Electronic
1176 Health Records and Medical Claims Data to Support Regulatory Decision-Making for Drug
1177 and Biological Products; 2021 Sep.
- 1178 FDA, United States. United States. Final Guidance: Best Practices for Conducting and
1179 Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data; 2013
1180 May.
- 1181 FDA, United States. Framework for FDA's Real-World Evidence Program. Silver Spring
1182 (US); 2018 Dec.

1183 FDA, United States. Guidance for Industry: Considerations for the Use of Real-World Data
1184 and Real World Evidence To Support Regulatory Decision-Making for Drug and Biological
1185 Products; 2023 Aug.

1186 FDA, United States. United States. United States Guidance for Industry: Data Standards for
1187 Drug and Biological Product Submissions Containing Real-World Data; 2023 Dec.

1188 FDA, United States. Guidance for Industry: Real-World Data: Assessing Registries to
1189 Support Regulatory Decision-Making for Drug and Biological Products; 2023 Dec.

1190 FDA, United States. Patient-Focused Drug Development Guidance Series for Enhancing the
1191 Incorporation of the Patient's Voice in Medical Product Development and Regulatory
1192 Decision Making; 2023 July.

1193 Health Canada, Canada. Elements of Real-World Data/Evidence Quality Throughout the
1194 Prescription Drug Product Life Cycle; 2019 Mar.

1195 ICH. ICH Reflection Paper Proposed ICH Guideline Work to Advance Patient Focused Drug
1196 Development; 2021 Jun.

1197 NMPA, China. Guidance for Real-World Data Used to Generate Real-World Evidence; 2021
1198 Apr.

1199 NMPA, China. Guidance on the Use of Real-World Evidence to Support Drug Development
1200 and Regulatory Decisions; 2020 Jan.

1201 NMPA, China. Guidance on Using of Real-World Study to Support Pediatric Drug
1202 Development and Regulatory Evaluation; 2020 Aug.

1203 MHLW/PMDA, Japan. Basic principles on the utilization of health information databases for
1204 Post-Marketing Surveillance of Medical Products; 2017 Jun.

1205 MHLW/PMDA, Japan. Guidelines for the Conduct of Pharmacoepidemiological Studies in
1206 Drug Safety Assessment with Medical Information Databases; 2014 Mar.

1207 World Health Organization. A Practical Handbook on the Pharmacovigilance of Medicines
1208 Used in the Treatment of Tuberculosis: Enhancing the Safety of the TB patient. Geneva (CH):
1209 WHO Press; 2012.

1210 www.fda.gov (Patient-Focused Drug Development Glossary 2018.)

1211 www.fda.gov FDA Patient-Focused Drug Development Guidance Series for Enhancing the
1212 Incorporation of the Patient's Voice in Medical Product Development and Regulatory
1213 Decision Making 2024.

1214 www.imi-prefer.eu (How to do a patient preference study.)

1215 www.reganudall.org (Real World Data: Assessing Electronic Health Records and Medical
1216 Claims Data to Support Regulatory Decision Making for Drug and Biological Products
1217 Guidance for Industry 2021 Nov.)

1218 **14 規制当局以外による関連するガイドライン**

- 1219 International Society for Pharmacoepidemiology. Guidelines for Good
1220 Pharmacoepidemiology Practices (GPP). Washington, DC (US): ISPE; 2015 Jun.
- 1221 The REporting of studies Conducted using Observational Routinely-collected health Data
1222 (RECORD) Statement.
- 1223 The REporting of studies Conducted using Observational Routinely-collected health Data
1224 Statement for Pharmacoepidemiology (RECORD-PE).
- 1225 “Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly work in
1226 Medical Journals,” established by the International Committee of Medical Journal Editors
1227 (ICMJE).

1228 **15 参考文献**

- 1229 1. Wang SV, Pottegard A, Crown W, Arlett P, Ashcroft DM, Benchimol EI, et al.
1230 HARmonized Protocol Template to Enhance Reproducibility of Hypothesis
1231 Evaluating Real-World Evidence Studies on Treatment Effects: A Good Practices
1232 Report of a Joint ISPE/ISPOR Task Force. *Pharmacoepidemiol Drug Saf.*
1233 2023;32(1):44-55.
- 1234 2. Gatto NM, Campbell UB, Rubinstein E, Jaksa A, Mattox P, Mo J, et al. The
1235 Structured Process to Identify Fit-For-Purpose Data: A Data Feasibility Assessment
1236 Framework. *Clin Pharmacol Ther.* 2022;111(1):122-34.
- 1237 3. Schneeweiss S, Rassen JA, Brown JS, Rothman KJ, Happe L, Arlett P, Dal Pan G,
1238 Goettsch W, Murk W, Wang SV. Graphical Depiction of Longitudinal Study Designs
1239 in Health Care Databases. *Ann Intern Med.* 2019 Mar 19;170(6):398-406).
- 1240 4. Hernan MA, Robins JM. Using Big Data to Emulate a Target Trial When a
1241 Randomized Trial Is Not Available. *Am J Epidemiol.* 2016;183(8):758-64.
- 1242 5. Wang SV, Pinheiro S, Hua W, Arlett P, Uyama Y, Berlin JA, Bartels DB, Kahler KH,
1243 Bessette LG, Schneeweiss S. STaRT-RWE: structured template for planning and
1244 reporting on the implementation of real world evidence studies. *BMJ.* 2021 Jan
1245 12;372:m4856. doi: 10.1136/bmj.m4856. PMID: 33436424; PMCID: PMC8489282.