

### 第3章 エビデンスに基づくレビュー方法

#### 定義とその範囲

このプロジェクトにあたり、カリフォルニア大学サンフランシスコ校とスタンフォード大学合同の UCSF-Stanford Evidence-based Practice Center (以下 EPC) は、*患者安全の実践の定義*を、「状態や行為において医療システムの提供によって生じる有害事象の確立を減少させるある種のプロセスまたは構造」とした。この定義に当てはまるような実践例は、オーダーエントリーシステム (CPOE) (6 章)、入院患者の血栓塞栓症予防 (31 章)、高齢の入院患者にみられる転倒予防対策 (26 章)、新しい教育技法である「クルーリソースマネジメント」(44 章) などが挙げられる。それに対して、医療を起因とする合併症というよりは、むしろその疾患に特有な、あるいは基礎疾患やその合併症 (急性心筋梗塞の治療でのアスピリンや プロロッカーの使用) に対する実践は、「患者安全の実践」には含めなかった。これらの問題についての議論は 1 章にて扱っている。

いくつかの事例では、患者安全とより一般的な医療の質の向上に対する対策とを区別することは難しい。*医療の質向上の実践*は、現在の医療の質のレベルが「危険」だとされる場合は患者安全の実践とみなすが、安全性を測る基準はしばしば変わり、数値で表わすのは困難であり、時間の経過によって変化するものである。例えば、心電図や X 線写真の解釈における精度のレベルにおいて、「適切レベル」や「安全レベル」とは何であるのか。少なくとも「適切」とされるところまでパフォーマンスを改善するため、患者安全実践についてはかなりの考察が必要であろう。というのも安全実践を行うことは手抜きのために生じる診断ミスをたくさん減少させているからである。他方では、我々は医療の質向上の実践としては最低限とされるレベル以上にパフォーマンスを改善することを主に意図して、実践を考察した。後者の例としては、乳房レントゲン撮影による乳房のスクリーニング検査の感度を改善するためのコンピューターアルゴリズムの利用がある[1]。

我々は入院患者や外来患者の橋渡し役として、急性期病院あるいは急性期治療に関わる実践を広く取り扱った。これは、大多数の患者安全の文献が急性期医療に関するものであるという事実と、より規制されている病院環境におけるシステムの変更は、より効果的に行えるであろうという考えを反映している。しかしながら、病院以外で行われる実践も、考察に含めた。例えば、褥瘡の予防 (27 章) を扱ったこの報告書は、病院のみならずナーシングホームにおいても適用可能である。

米国政府の医療研究・品質局 (The Agency for Health Care Research and Quality : 以下 AHRQ) (この報告書を委託されている団体) による本報告書がカバーすべき範囲を EPC チームは入手した。加えて EPC チームは、2001 年 1 月 26 日に行われた National Quality Forum (NQF)

Safe Practices Committee の公開講座に参加した。NQF は 1999 年に消費者、購買者、提供者、医療保険会社、医療サービス研究団体が集まり、質の向上に対する国家的対策をつくるために組織された団体である。安全実践委員会のメンバーが EPC チームと共同で行った仕事が発展しこの報告書の形をなしてきた。

### 領域の構成

患者安全の可能性に関する評価・同定を促進するために、プロジェクトの内容を様々な領域に区分した。いくつかの領域は薬剤有害事象、院内感染、外科手術合併症のような“内容区域”を扱い、また他の領域では患者安全の向上（情報技術、人的要因研究、組織理論）が期待されるアプローチを含む広範な（医学以外の）専門分野における実践を対象とする。領域は文献から一般的に解釈されたものであり、可能な限り包括的にした。リストは、完全性という点で患者安全の専門家、臨床研究者、AHRQ、NQF 安全実践委員会のレビューを受けた。それぞれの領域では、主題となっている事項の専門家、もしくはエビデンスに基づくレビューや技術面での批判的吟味に精通している者を著者や共編者として選んだ。この報告書の著者は、米国の主要な学術センターに所属しており、その名前を本書の最後に列挙している。

### 評価を行うための安全実践の同定

UCSF-Stanford Evidence-based Practice Center の調整チーム（編集委員会）は、安全実践の同定し、評価を行う際の検索方針に関する全般的な指示（表 3.1）を著者らに伝えた。必要に応じて編集委員会はガイドラインを加えたり、文献の追加検索を行った。

表 3.1 調整チームが著者に推奨する検索方針

- |  |
|--|
| <p>a) 電子文献データベースの使用。MEDLINE や Cochrane Library を用いて全ての検索を行うこと。トピックについては Cumulative Index to Nursing &amp; Allied Health (CINAHL) や PsycLit (PsycINFO)、ISI 社の Science Citation Index Expanded・Social Sciences Citation Index・Arts &amp; Humanities Citation Index、INSPEC (physics, electronics and computing)、そして ABI/INFORM (business, management, finance, and economics) による検索も必要となる。</p> <p>b) 検索された記事関連書籍や、主要雑誌の目次を手作業で検索する。</p> <p>c) 官報（Grey literature）や多くのトピックについて会議録、施設報告、博士論文、製造メーカーの報告書といった関連資料のレビューが必要である。</p> <p>d) 専門家やその分野の職員に相談する。</p> |
|--|

本報告書の早期の刊行を AHRQ が要求していたので、英語以外の言語による論文の検索や、EMBASE による検索は実施しなかった。著者チームの裁量で、重要なトピックに関する英語以外の論文を含めた。編集委員会はデータベースの検索の際に発表年の制限は加えなかった。このプロジェクトで特に重要なのは、患者安全のトピックに関連する系統的なレビューを行い、確認することである。系統的レビューを行うために伝達した方針は、全ての情報が入手可能なわけではないが、適切に MEDLINE など（OVID, Silver Platter）のデータベースを用いて検索を行うことである。要望があれば利用可能な検索手段として、National Library of Medicine の無料データベース検索サービスである PubMed も提供した[4]。編集委員会は特定のトピックについては独自に書籍データベースや他の参考資料を検索した[5]。同時に、EPC は NQF と協力して NQF 会員からエビデンスに基づく実践について情報を求め、レビューを行う追加文献を提供してくれた本プロジェクトの理事会に助言を求めた。

#### 実践の採用除外基準

EPC は正当に評価された安全実践を選択するための基準を設置した。基準はさまざまな状況や手順による実践に適用可能であり、実践の有効性や効果に関するエビデンスが入手可能なものとした。

**表 3.2 実践の採用するか除外するかの基準**

##### 採用基準

1. 実践が病院、および入院患者、外来患者に適用できるものであり、かつ幅広い範囲で医療ケアの状況や手順に適用できること。
2. 安全実践のエビデンスは、少なくともひとつはレベル 3 よりも高い研究デザインのもが含まれ、かつアウトカムの尺度がレベル 2 であること。診断や治療上の介入に特に関係しない実践については、レベル 3 のアウトカムが適当である。（レベルの定義については表 3.3 を参照。）

##### 除外基準

1. 実践について上記の方法の基準に合う研究が存在しない。

病院外で行われたり、特殊な状況の患者、もしくは特定の医療処理を受ける患者についてのみ研究された実践については、その実践が病院内やさまざまな状態や手順にも適用できると著者と編集委員会が判断した場合は、それを含めることにした。潜在的な可能性のある安全実践の数を増やすため、医療以外の分野からも実践を取り入れ、診療や治療など

によほどのこと関係しない実践でない限り、患者安全の測定があまり厳密ではない研究によるエビデンスも採用した。これらの基準により、チームワークトレーニング（第 44 章）や情報伝達改善方法（第 42 章）といった領域も含められた。

それぞれの実践の有効性や効果についてのエビデンスのレベルは、研究デザイン（表 3.3）と研究アウトカム（表 3.4）という点から評価された。編集委員会はエビデンスを評価するための既存の枠組みを修正し[6-8]、エビデンスの統合に関するその他多数の情報源から得られた勧告を含めて次の階層を作った[9-25]。

**表 3.3 研究デザインの階層**

レベル 1	<u>ランダム化比較試験</u> - 代替配分といったランダム化に準じるプロセスも含む
レベル 2	<u>非ランダム化比較試験</u> - 適格基準とアウトカム尺度が予め定めた前向き（事前計画）研究
レベル 3	<u>比較対照群を用いた観察研究</u> - 後ろ向き研究、時系列が中断（介入に伴う傾向が変化）している研究、ケースコントロール研究、比較対照群を用いたコホート研究、交絡変数の影響を調整した医療サービス研究も含む。
レベル 4	<u>比較対照群を用いない観察研究</u> - （例：比較対照群を用いないコホート研究やケースシリーズなど）
*系統的レビューとメタ分析はレビューに含める最高レベルの研究デザインとし、「A」を加えた（具体的には少なくともひとつのランダム化比較試験を含む系統的レビューは「レベル 1A」とした）。	

**表 3.4 アウトカム尺度の階層**

レベル 1	<u>臨床アウトカム</u> - 罹患、死亡、有害事象
レベル 2	<u>代替的アウトカム</u> - 観察されたエラー、臨床的なアウトカムの利害（通常は有害事象）に密接なつながりをもつ中間的アウトカム（検査結果）
レベル 3	<u>その他の測定可能な変数で、対象とする安全性に確実なつながりがない、もしくは間接的なつながりしかないもの</u> - （例：教育的介入後の事前/事後テスト、様々な実験的状況における操作者の自己報告）
レベル 4	<u>医療エラーや有害事象の減少に関連するアウトカムが存在しないもの</u> - （例：患者満足度が測定された唯一のアウトカムである研究、エラーを検出するアプローチについて記した論文ではあるが、アウトカムが測定されていないと報告しているもの）

アウトカム尺度のこれらの暗黙の階層は、代替的アウトカムや中間的アウトカム(レベル2)が臨床アウトカム(レベル1)の利害関係に確実に関係する[26]。患者安全に関連するアウトカムではあるが、罹患や死亡に伴うアウトカムでないものはレベル3に分類される。

#### *EPC 基準の例外*

いくつかの安全実践は有効性や効果に関するエビデンスが乏しいため、EPCの基準には合致しなかったが、表面的妥当性があれば報告書には含めた。(詳しい情報を周知している読者は表面的妥当性があるかどうか評価が行われることを当然求めるだろう。第1章を参照。)これらの実践のレビューは、医学や以外以外の分野から選りすぐられたエビデンスの質を明らかにするだろう。

#### **安全実践の評価**

それぞれの実践について、著者は次の情報について文献を調べるよう指示された。

- ・ 実践の対象とする問題の頻度
- ・ 実践の対象とする問題の重大性
- ・ 実践の現在の状況
- ・ 実践の有効性や効果に関するエビデンス
- ・ 実践の潜在的な有害性
- ・ 入手可能であればコストに関するデータ
- ・ 実施上の問題

特に主題やエビデンスの質が幅広く様々であれば、これらの要素は全体の章にできるだけまんべんなく盛り込むよう努力した。それぞれの実践に関する資料が多いことが期待されるし、事実そのように期待したので、編集委員会は実践の有効性や効果に関するエビデンスの報告や検索のためのプロトコルと共に、それぞれの要素に何が要求されているかという一般的な指針を呈示した。

プロトコルは、検索、研究への採用基準、研究の要約における要素、そしてそれぞれの研究から報告された情報の指針などを概説した。著者は実践を確認するために検索を行って論文をレビューし、よりよい研究デザインで実施された実践を収集することが求められた。トピックによってはさらに重点的に検索が行われた。研究に採用するかどうかの基準は、研究デザインに直接関連する。著者は、エビデンスがあるレベルの質を満たす研究デザインであるか、もしくはレビューされるのに必要な次のレベルのエビデンスであるかどうかを判断することが求められた。デザインの質が低い研究を除外することを正当化する

ためには、その根拠となる少なくとも二つの十分な質の研究があることを編集委員会は示した。たとえば、2つの適切な臨床試験（レベル1のデザイン）が含まれていることは、同じトピックについて前向き、非ランダム化試験（レベル2）から入手したエビデンスを除外するために必要となる。

編集委員会は研究デザインに基づいて、研究の採用基準に合致する論文を要約する際の指示を示した。それぞれの研究について、必須となる10の要約方法（表3.5）を明確に示した。著者は要約に関するそれぞれの必須要素について詳細に説明を受け、さらに研究デザイン要約の3タイプの例を渡された（レベル1A、1、3、レベル2は集めた情報がレベル1と同じであったため含めなかった）。研究チームは特定の主題領域に関連するどのような追加情報も要約するよう奨励された。

**表 3.5 10の必須要約要素**

<p>1. 「AMA Manual of Style」に準拠した書籍情報：タイトル、著者、出版日、情報源</p> <p>2. 研究デザインのレベル（効果の情報を示す研究はレベル1-3、関連追加情報が必要な場合はレベル4）とともに次に説明する資料があること。</p> <p><i>記述子</i></p> <p><u>レベル1 系統的レビューのみ</u></p> <p>(a) 記述方法が示された情報源を確認することができ、論文を検索できること。</p> <p>(b) 論文の採用・除外基準について説明している：はい/いいえ</p> <p>(c) 文献に含める研究の範囲</p> <p><u>レベル1か2の研究デザイン（系統的レビューなし）</u></p> <p>(a) 盲検化：盲検化されている、盲検化が不明、盲検化されていない</p> <p>(b) ベースライン時点における群間の比較可能性を記述。すなわち、潜在的な交絡因子の分布はベースライン時点で等しかったか？もし等しくなかったなら、どの交絡因子が等しくなかったのか？</p> <p>(c) 全体的な追跡脱落者：全研究集団のうち、追跡脱落者の割合。</p> <p><u>レベル3の研究デザイン（系統的レビューなし）</u></p> <p>(a) 研究デザインの記述（例：ケースコントロール、時系列）</p> <p>(b) ベースライン時点における群間の比較可能性に関する記述。すなわち潜在的な交絡因子の分布はベースライン時点で等しかったか？もし等しくなかったなら、どの交絡因子が等しくなかったのか？</p> <p>(c) 潜在的な交絡因子を調整した分析である：はい/いいえ。はいに該当する場合</p>
---

は、どの交絡因子を調整したか？

3. 介入の記述（可能な限り具体的に）
4. 研究の対象集団の数と設定の記述
5. 関連するアウトカム尺度レベル（例：レベル4）の記述
6. 関連するアウトカム尺度記述
7. 主要結果：効果の大きさと信頼区間
8. 実践の予期しない有害な（便益のある）影響に関する情報
9. 実践のコストの情報
10. 実践の実施に関するの情報（実践が役に立つかどうか、どのように実施するのかという情報。例：実施についての既知の障壁）

我々は採用されたそれぞれの研究（例：研究デザイン、対象集団の数と設定、介入の詳細、結果）について重要な要素を文章や表で示した。加えて、結果の解釈が不自然である場合、研究の弱点とバイアスを明らかにするように著者に求めた。著者は正式には報告書のための研究の範囲を超えてエビデンスを組み合わせたり、統合したりすること（例：メタ分析を実施するなど）は求められていない。

### レビューのプロセス

著者は2段階で編集委員会に作業結果を提出する。最初の段階（「安全実践の評価の確認」）では、著者が任命されてからおおよそ6週間後に提出され、著者チームはレビューを行う検索方針、引用、患者安全の実践の予備リストが渡された。次の段階では（「安全実践の評価」）、任命されてからおおよそ12週間後の期限で、著者はそれぞれのトピックについてその章の草案の提出、要約書式への記入、見直しの繰り返し、校正をおこない、最後にその章が完成する。

#### 安全実践の評価の確認

編集委員会と理事会はドメインのリストと実践とのギャップを埋めるためにレビューした。さらに、編集委員会は最終的に著者が提出した除外された実践のリストとともに、除外が正当だとする理由についてもレビューを行った（例：不十分な研究デザイン、不十分なアウトカム、単一の疾病のプロセスに特有な実践である）。実践を採用するか、除外するか基準に合致しているかどうか意見が異なる場合は、編集者は著者と相談した後、最終的な判断を下した。報告書の目次にある最終的な実践のリストは、AHRQとNQF安全実践

委員会にコメントをもらうため回覧に供した。

#### 安全実践の評価

編集者のチームによって各章はレビューされ（EPC の調整チームの編集委員会と我々の編集長）、質問は著者に伝えられ、多くの場合さらなる分析や、結果や結論の拡大が求められた。各章が完成したら、報告書全体は、重複する題材を削除し、安全に関する実践に重点が置かれたままであるかを確認するために編集された。レビューのプロセスのほぼ終わり近くで、コメントをもらうために理事会に配布した。いったん内容が仕上がったら、編集者は 56 章に記した方法論を使い実践を分析し、ランク付けした。これらの要約とランク付けの結果は報告書の 56 章以降に示した。