

文献チェック用 ワークシートチェックポイントとその解説

a)治療の文献のチェックポイント

対象となる研究デザイン¹⁰：ランダム化比較試験（なければ擬似ランダム化比較試験や、非ランダム化比較試験など）

対象患者： _____
 検討した治療内容： _____
 比較した治療内容： _____
 評価基準とした転帰： _____

文献の結果は信頼できるか

ランダム割付による比較試験であるか¹¹。

（はい、いいえ、不明）

研究対象患者のすべてが評価の対象となっているか。治療中の脱落症例はないか¹²。最初の割り付けられた群のまま評価されているか（intention to treat analysisか）¹³。

（はい、いいえ、不明）→いいえの場合、それは結果をくつがえしうるほど重大か
 効果判定にあたっては、患者の治療内容を知らされないまま治療評価をしているか¹⁴。

（はい、いいえ、不明）

試験の最初の段階では、両群の背景に差はなかったか¹⁵

（はい、いいえ、不明）→いいえの場合、それは結果をくつがえしうるほど重大か

検討対象となった治療内容以外に、両群の治療内容に差がないか¹⁶。

（はい、いいえ、不明）→いいえの場合、それは結果をくつがえしうるほど重大か

結果はどのようなものか

まず、以下の表を埋めて考えよう（転帰とは評価基準となった出来事を指す。研究で複数の転帰が検討されていれば、患者にとって重要なものを重視するのが原則。）

	転帰発生	転帰発生せず
検討した治療		
標準的治療		

治療群と対照群とで転帰の差はどの程度のものであったか

（改善する、差はない、悪化する、不明）

その結果からの予測はどの程度正確と考えられるか¹⁷

（正確、まあまあ正確、あまり正確でない、不正確）

この文献の内容は、今の状況に当てはめて良いか

研究対象は、患者と同じ様な患者群か。（研究対象患者に、自分が想定している患者は含まれるか）

（はい、いいえ、不明）

治療内容は、現状で行なえるものか。

（はい、いいえ、不明）

治療内容は、患者にとって容認できることか。リスクやコストを考慮する。

（はい、いいえ、不明）

¹⁰ このチェックポイントは、あくまでもランダム化比較試験を念頭に置いたチェックポイントである。ただ、ランダム化比較試験がない場合には、最初の質問を「いいえ」として用いることになる。

¹¹ ランダム化の手順が重要である。割付表・乱数表が現場から離されているか、事前にわからないような方法であったかなども確認する。「ランダム化は現場から隠されていたか」というチェックポイントを加えることがある。

¹² 脱落症例とは、追跡中に経過がわからなくなった対象者を指す。予定された治療が行えなくても、その経過が追跡され予後の検討に加えられていれば脱落症例ではない。脱落症例は少ないほど妥当性が高い。原因のわからない脱落や、治療効果や経過に関連する脱落は結果の妥当性を脅かす。転帰の発生率が低い場合には、わずかな脱落症例の転帰によって結果が覆ってしまうので、特に問題になる。

¹³ Intention to treat analysisとは、最初の割付に従ってそのまま集計する方法である。これによりランダム化による2群間の均一さが守られる。また、現実には治療を行う予定でも副作用や諸事情で続けられないこともある。このような場合もそのまま検討することで、現場での判断そのものをより厳しく評価したことになる。ITT分析と略される。

¹⁴ 治療効果を検討する評価者が治療内容から知っているのは効果の判定に影響を与える。特に、症状や臨床所見などの評価者の主観が入りやすいものが問題になる。死亡などでは評価者の主観は入りにくく、問題は少ない。

¹⁵ ランダム化を行っていても、対象者が少ない場合には両群の背景がそろわず予後に大きな影響を与える因子に差が生じる場合がある。このような偏りがなかったかどうか、両群の患者背景を比較して検討する。

¹⁶ 治療に当たっている医療チームなどが割付を知っていると、他の治療内容が大きく偏る場合がある。医療チームが盲検化され割付内容を知らされていないければ問題は少ない。盲検化ができない場合には、治療手順を定めたプロトコールなどを作る、併用療法について細かく記録を取り後に差がなかったか検討する、などの対策が取られる。

¹⁷ ここでの評価項目は便宜的であり、通常は治療効果の指標の信頼区間を示すとよい。

b) 予後に関する文献のポイント 対象となる研究デザイン：コホート研究、ランダム化比較試験などもコホート研究の一種である。

対象者： _____

予後の評価の基準とした障害・疾患・死亡など： _____
 (研究で検討された予後に関連する危険因子： _____)

結果は信頼できるか

対象となった患者はその疾患の同等の病期（重症度、進行度など）から得られているか¹⁸

(はい、いいえ、不明)

追跡は十分長く十分完全に行なわれたか¹⁹

(はい、いいえ、不明)

転帰は客観的でバイアスを避けたものが用いられたか²⁰

(はい、いいえ、不明)

重要な予後決定因子に関して補正を行ない検討されているか²¹

(はい、いいえ、不明)

結果はどのようなものか

経過に伴う転帰はどのようなものか（なるべく実際の頻度・確率を見る。複数群で比べている場合には、予後の差を見る）

重要な事故(転帰)発生率は？

予後に関する予測はどの程度正確か

重要な事故(転帰)発生率の信頼区間は？

結果は自分のそして自分の患者の役に立つか

対象患者は自分の患者と似ているか

(はい、いいえ、不明)

その結果が治療の選択に関わるか

(はい、いいえ、不明)

その結果を告げることが患者の役に立つか

(はい、いいえ、不明)

¹⁸ たとえば肺ガンを想定した場合、そのガンの進行度（TNM分類など）や病理診断（扁平上皮癌、腺癌、小細胞癌など）によって予後は大きく異なる。これらを考慮せずに「肺ガン」としてまとめただけの予後の評価では、実際に具体的な患者を念頭に置いて予後を予測する上では使いにくい情報が得られない。

¹⁹ どれほどの追跡が十分かは、疾患によって異なる。癌であれば数年にわたる追跡が必要であるうし、咽頭炎であれば数週間でよいかもしれない。また、どれだけの追跡率を十分とするかは、その疾患の経過による。ほとんど死亡などの転帰が発生しない場合には、わずかな脱落症例によって大きく予測がずれてしまう危険性があるので、脱落の少ない追跡結果が必要になる。

²⁰ 死亡などの客観的な転帰であれば問題はないが、重症化したかどうかという転帰であれば、広く用いられ妥当性や再現性も検討された指標を用いて客観性を確保しなければ、得られた定量的な結果も汎用性の高い情報とはならない。

²¹ たとえば、痴呆患者の予後をアルツハイマー型痴呆と脳血管性痴呆、その他の神経疾患による二次的痴呆とで検討した場合、それぞれの年齢や性別によって補正を行わなければ、それぞれの疾患による予後への影響がどれほどかわりにくくなる。予後に影響を与えらる重要な要因に関しては、統計学的な処理を行ってうち消して検討する。この方法を補正という。

c) 診断に関する文献のチェックポイント

対象となる研究デザイン：横断研究（確定診断までに時間がかかるような場合には、経過調査が必要）

対象患者： _____

検討した検査： _____

評価基準とした確定診断： _____

結果は信頼できるか

標準的な、診断を確定する検査（gold standard）²²と、独立にしかも盲目化されて比較されているか²³

（はい、いいえ、不明）

実際にその検査の対象となる患者群において検討されているか²⁴

（はい、いいえ、不明）

研究の対象となった検査結果に関わらず、標準的な検査が行なわれているか²⁵

（はい、いいえ、不明）

検査方法は実行可能なように明確に記載されているか

（はい、いいえ、不明）

結果はどのようなものか

感度、特異度、尤度比はどうか

（診断に役立つ、役立つかない、不明）

以下の表を埋めて考えよう

	疾患あり	疾患なし
検査陽性		
検査陰性		

追加のポイント²⁶

その検査が陽性の時、確定診断ができるか

（はい、いいえ、不明）

その検査が陰性の時、除外診断ができるか

（はい、いいえ、不明）

その結果は自分の患者にあてはまるか

自分の施設、状況で行えるか

（行える、行えない、不明）

結果は自分の患者に当てはまるか

（当てはまる、当てはまらない、不明）

その検査は自分の医療行変えるか²⁷

（変える、変えない、不明）

その検査によって患者に利益があるか²⁸

（利益がある、利益はない、不明）

²²確定診断のための検査の例：狭心症は冠動脈造影。細菌性髄膜炎は髄液検査。乳ガンであれば生検や手術による病理診断となる。しかし機能的な評価が重要である場合には疾患（痴呆症など）では、一定の診断基準を設定しそれを満たすかどうかで判断する。また、侵襲性などの問題から確定診断のための検査が困難である場合（ウイルス性脳炎に対する脳生検、急性肺血栓塞栓症での肺動脈造影）もある。可能な限り、確定診断結果と検討対象の検査結果を比較する必要がある。

²³診断の研究でも目隠しは重要である。確定診断結果と、検査結果とはお互いの結果を知らずに独立して判定されなければならない。たとえば、胸部レントゲン上肺炎が疑われることをした上で肺の聴診をすれば、より異常音に気づきやすくなるし、逆に臨床所見を知ってレントゲン写真を読影すると曖昧な肺陰影を肺炎としたりしなかったりするかもしれない。

²⁴急性心筋梗塞の検査を検討した研究で、検討対象が救急外来で心筋梗塞と確定診断された患者と、成人ボランティアとの比較では、あまり現場での意味はない。このような比較のために検査を行うことは考えにくい。同じ研究でも、救急外来を受診した胸部症状を訴えて患者群で検討していれば、実際に検査の対象となる患者群で検討されたと判定される。

²⁵ある基準を満たした患者全員について、検査結果と確定診断結果の両方を知る必要がある。検査が陰性であればすべて確定診断も陰性とする、検査が陰性なのに実は疾患があるという見落とし症例がないことになる。このような見落としがないように工夫していなければ、偽陰性をすべて見落としとして検討することになる。

²⁶これは、現場での利用価値を考えるための付け加えたポイントである。もし、その検査が確定診断や除外診断に役立つのなら、利用価値は高い。

²⁷その検査結果を知ったところで、医療側に何ができるかを確認しておく。

²⁸診断をしても、特に患者の経過に大きな影響はなく、患者にとっては心配と疑念しか与えない、ということであれば、患者にとって意味は小さい。また、その検査によってより侵襲性の高い治療や検査を患者に行うことになり、かえって患者の予後に悪影響を与える危険性はないだろうか。検査の有効性の判定には十分注意が必要である。

d) 病因・副作用に関する文献

対象となる研究デザイン：症例対照研究、コホート研究、ランダム化比較試験

対象患者： _____

検討した危険因子： _____

危険の評価の基準とした障害・疾患： _____

結果は信頼できるか

比べている患者群は、その副作用の原因と考えられる要因以外は同様であったか²⁹

(はい、いいえ、不明)

その要因に関しても、転帰予後に関しても両群で同じように評価されているか³⁰

(はい、いいえ、不明)

追跡は十分長く十分完全に行なわれたか³¹

(はい、いいえ、不明)

時間的前後関係は正しいか³²

(はい、いいえ、不明)

量反応関係がなりたつか³³

(はい、いいえ、不明)

結果はどのようなものか

要因と結果との間に強い関連があるか

(強い、弱い、明らかでない、不明)

危険性はどれほど正確に予測されるか

(かなり正確、あまり正確でない、不確か、不明)

結果は自分のそして自分の患者の役に立つか

その結果は自分の患者にあてはまるか

(当てはまる、当てはまらない、不明)

危険はどの程度のものか³⁴

(重大、あまり重大でない、軽微、不明)

その要因を排除しようとするべきか³⁵

(するべき、した方がよい、しなくてもよい、しない、不明)

²⁹ 症例対照研究の場合、結果と思われる疾患・病態・副作用のあるなしで対象者を選び、要因を比較する。このときに、検討しようとする要因以外は似た対照者を選ぶ手順を行う。これをマッチングという。要因の違いの影響を検討しやすくする。検討する研究が、コホート研究であれば要因の違いを補正する手順を取ったかどうか、ランダム化比較試験であれば両群が似ていたかどうかを検討する。

³⁰ 症例対照研究の場合には、その要因の調査にあたって両群とも同じような手順を踏む必要がある。ふつう、合併症と思われる疾患に罹患した方が、疑わしい要因を思い出しやすい(思い出しバイアス)。客観的な指標を用いたり、詳細な裏付け調査を行うなどの工夫をする必要がある。

³¹ ここでも、検討対象となっている要因や副作用、その頻度などを考慮して判断する。

³² 時間的前後関係とは、その結果と思われる転帰の前に原因と思われる要因が関与したという関係である。慢性的な疾患・副作用を検討している場合、いつ発症したかがわかりにくいことがある。このような場合には要因に曝露されてからどれくらいかかって疾患が成立するかの判定が難しい場合もあるので注意が必要である。

³³ 量反応関係とは、危険因子が大きいほど結果も大きいという関係である。これが見られると、因果関係をより強く示唆する結果であると判断される。ただ、アレルギー反応などでは量の多い少ないよりも、曝露されたかどうかが重要な意味を持つので量反応関係は見られない。これが見られないからといって因果関係が否定されるわけではない。

³⁴ ここでの危険とはただ単に患者の視点からだけでなく、患者家族や社会、医療従事者側など、いろいろな立場から検討する。

³⁵ 具体的な排除方法を考慮して判断する必要がある。その要因が簡単に取り除けたり、他に代替手段があれば排除は容易になる。疑われる要因が薬剤であった場合、その薬剤の有効性・重要性和その副作用のリスクを比較する必要がある。

e) システマティックレビュー・メタアナリシス

検討の対象となる研究デザイン：治療に関するシステマティックレビュー

診断など治療以外のシステマティックレビューの場合には、後半のチェックポイントを診断や予後などのものを用いる

対象患者：_

検討した治療内容：_____

比較した治療内容：_____

評価基準とした転帰：_____

結果は信頼できるか

そのレビューは焦点を絞った臨床からの課題をもとにしているか³⁶

(はい、いいえ、不明)

レビューの対象とする論文の基準は適切か³⁷

(はい、いいえ、不明)

重要な研究が見逃されていないようにされているか³⁸

(はい、いいえ、不明)

対象となった論文の信頼度に関する批判的吟味が行なわれているか

(はい、いいえ、不明)

それぞれの研究の結果は均一か³⁹

(はい、いいえ、不明)

結果はどのようなものか

レビューの全体の結果は何か⁴⁰

その結果はどれほど正確と考えられるか

結果は自分のそして自分の患者の役に立つか

その結果は自分の患者にあてはまるか

(当てはまる、当てはまらない、不明)

すべての重要な転帰が考慮されているか

(はい、いいえ、不明)

その益は、害やコストに見合ったものか

(はい、いいえ、不明)

³⁶ 過去の研究をまとめるときに、ある程度焦点を絞って作業をはじめなければならない。そのためには、対象患者、検討した治療、比較した治療、評価基準が明確に定義されている必要がある。最初の曖昧さが大きければ、後で足し合わせることができないような研究が集まることになる。

³⁷ ある程度の質の基準を満たしていなければ、足し合わせる価値はない。治療に関する系統的総説では、通常ランダム化比較試験が選ばれる。時には擬似ランダム化比較試験まで対象とすることがある。診断に関するシステマティックレビューでは、検査と確定診断が別々に独立して行われたものが優先される。予後であれば、事前に登録基準が明記されたコホート研究が優先される。

³⁸ 文献検索としては複数の電子データベースを探し、専門家・研究者にも問い合わせ、対象となった論文の参考文献まで検討対象にしたかどうかを確認する。学会の抄録集をハンドサーチすることもある。検討した文献を英語に限った場合には、結果の偏りが生じる危険性が増す。できるだけあらゆる言語を対象にしている方がよい。

³⁹ メタアナリシスを行うためにそれぞれの研究に不均一性 (heterogeneity) の有無の検討が行われる。この検討によって、複数の研究結果の均一性 (homogeneity) が認められるということが確認されれば、研究結果の集計結果は受け入れやすい。もし、不均一性があるとされた場合には、すぐメタアナリシスを行わずまずその原因を検討する。対象患者や治療方法、検査方法などの違いが不均一性の原因であることがある。また、研究そのものの信頼性や研究の行われたに時期や背景によりばらついていることがある。このような検討から、治療効果や検査の解釈に重要な要因や傾向が見つけられることがある。このような検討を行った上でも、結果の集計が困難な場合には、バラツキの影響をそのまま受け入れる解析方法 (ランダムイフェクトモデル) をとることが多い。

⁴⁰ メタアナリシスを行った場合には、オッズ比やリスク比、検査であればROC曲線といった統合しやすい指標で結果が示される。予後であれば単に生存率とその信頼区間という場合もある。何がどう集計されているか注意すること。特に治療の場合には、3つ以上の治療法を検討し、複数の評価基準での結果をまとめ、いくつもの指標が示される場合がある。それぞれの指標が何と何をどのような基準で比較しているのかに注意して検討する必要がある。