



# 臨床研究と疫学研究における論文の 質を高めるための国際動向： 人を対象とした研究デザインの エビデンス・グレーディング

上岡洋晴\*・津谷喜一郎\*\*・川野 因\*\*\*・武藤芳照\*\*\*\*・  
塩澤信良\*\*\*\*\*・宮本義久\*\*\*\*・本多卓也\*\*\*\*

(平成 19 年 10 月 5 日受付/平成 20 年 1 月 18 日受理)

要約：本研究は、人を対象とした治療や健康増進の効果を明らかにする研究のエビデンスを示す方法をレビューすることを目的とした。疫学や臨床研究においては、研究デザインによって、エビデンスのグレーディング (grading : 格付け) が行われている。各種の研究デザインにおける論文を評価したり、論文の質を高めるためにチェックリストや声明が開発されている。また、人を対象とした介入研究を実施する場合には、パブリケーション・バイアスの減少、研究の倫理、臨床試験参加の推進につながるため、事前の臨床試験登録 (clinical trial registry : CTR) が必要である。世界的な 11 の医学雑誌からなる「国際医学雑誌編集者委員会 (International Committee of Medical Journal Editors : ICMJE)」は、臨床試験の登録 (clinical trial registry : CTR) を論文投稿の条件としている。疫学研究や臨床研究を実施する場合には、こうした動向を理解する必要がある。

キーワード：エビデンス・グレーディング, 疫学研究, 臨床研究, 研究デザイン, チェックリスト, 声明

## 1. 緒 言

1990 年代後半から日本においても「エビデンスに基づいた医療 (evidence based medicine : EBM)」や「エビデンスに基づいた健康政策 (evidence based health policy : EBHP)」など、「エビデンスに基づいた」という用語が頻繁に用いられるようになってきた。

人を対象とした医学あるいは関連領域の研究としては、臨床研究・疫学研究が主である。これらの研究領域では、エビデンスのグレーディング (grading : 格付け) が行われるとともに、論文の質を高めるために最低限必要な情報あるいは実施すべき分析方法を記載することが求められるようになってきた。これが、各種のチェックリストや声明である。学術雑誌の査読者や編集者、レビューアに限らず、研究者は、研究の質を高める研究やそのコンセンサスが得られている国際動向を理解しておく必要がある。

本論では、「evidence」(エビデンス) の定義を「科学的根拠」とし、人を対象とした治療や健康増進の効果を明らかにする研究のエビデンスを示す方法をレビューすること

を目的とした。

## 2. エビデンス・グレーディング

エビデンスの構築と整理のために、研究デザインによってエビデンス・グレーディングがなされるようになった (表 1)。「図解 国民衛生の動向」にも、このことが明記されている<sup>1)</sup>。常識的になりつつあるエビデンス・グレーディングの考え方は、ヒトゲノムや分子生物学、生理学などの基礎医学ではなく、生物統計学が用いられる疫学研究や臨床研究の研究デザインが対象となる。したがって、理学療法、音楽療法、動物療法、森林療法、運動療法、食事 (栄養) 介入などが、心身に及ぼす効果 (治療・健康増進) を評価する研究の場合には、研究デザインの計画段階において、この格付けを見越して実施することが重要である。

エビデンスのグレードで、最も真実を示す可能性が高いと位置づけられるのが、ランダム化比較試験 (RCT) の「システマティック・レビュー (メタ・アナリシスを含む) : SR」であり、ついで「RCT」、「非ランダム化比較試験 (nRCT)」、「観察研究 (OS)」と続き、最下位が「患者データに基づか

\* 東京農業大学地域環境科学部身体教育学研究室

\*\* 東京大学大学院薬学系研究科医薬政策学講座

\*\*\* 東京農業大学応用生物科学部公衆栄養学研究室

\*\*\*\* 東京大学大学院教育学研究科身体教育学講座

\*\*\*\*\* 岡山大学大学院医歯薬学総合研究科長寿社会医学講座

ない専門委員会や専門家個人の意見」となっている。例えば、表1において、「Ⅰの研究とⅢの研究で相反する結果が出た場合には、Ⅰの結果の方が真実である可能性が高い」として採択することになる。RCTが高いグレーディングである理由は、「(1) 評価したい因子以外の結果に影響を及ぼす潜在的な因子である交絡因子の影響を少なくし、(2) データのバラツキの大きさを群間で均一化する、という重要な役目<sup>2)</sup>があり、「疫学で利用できる科学的に最も厳密な仮説検証の方法とされている<sup>3)</sup>」ためである。

エビデンスについては、図1のように「つくる」「つたえる」「つかう」の3つの局面がある<sup>4)</sup>。現場での指導や臨床は、「エビデンスをつかう」の部分に該当し、臨床研究や疫学研究で得られた良好、つまり「効果が期待できる介入」方法を現場で役立てることとなる。他方の「つくる」と「つたえる」は、換言すれば「研究すること」と「良質な研究結果を正確にまとめること (SR)」に相当する。

### 3. 総説とシステマティック・レビュー

論文の種類として、「総説」では、自然科学分野で用いられている従来からの記述的レビュー (narrative review : NR) と2000年以降に急増してきたシステマティック・レビュー (systematic review : SR)<sup>5)</sup>があり、両者は完全に区別される。

NRは、論文の収集・採用時において、主観的な選択や著者の考えを支持する論文を選択的に採用する傾向が否定できないことや、エビデンスのグレードの低い論文まで議論の対象とするなどの問題点が指摘されている。さらに、「選ばれた文献の解析法は、カテゴリーごとに支持され

る論文数を数え、最も投票数が多いカテゴリーを選ぶといった非科学的な方法であり、各論文の研究デザイン、効果の大きさ、サンプルサイズを無視している。したがって、このレビューの結論は、別のレビューの結論としばしば矛盾を導き出してしま<sup>6)</sup>という問題点も指摘されている。

一方、SRは、研究テーマや選択基準などを明確に規定し、対象論文を選択的・網羅的に収集する方法である。論文の収集については、あるデータベースを用いて、適格基準 (除外基準を含む) に従って検索すれば、世界中の誰が行っても同一の論文がヒットすることになる。さらに、SRは最終的にすべての研究の結果を統合するメタ・アナリシスを行い、批判的吟味とともに一般化可能性 (外的妥当性) や全体のエビデンスを示すことができるところに大きな特徴がある。

エビデンスのグレードにおいて、最上位とされるRCTのSRは、とくに医学に関連する研究として、「コクラン共同計画<sup>7)</sup>」が有名である。エビデンスを構築するには、RCTのSRが重要であることに賛同した研究者らの集まりとして、1992年に発足し、独自にSRである「コクラン・レビュー」を発表している。「1) すべてのRCTから、2) 良いものだけを、3) まとめて、4) 遅れなく、5) 必要な人に届けること<sup>4)</sup>」を目標とし、SRの重要性を明確にした提唱者でもあるイギリスの医学者 Archiebald L Cochrane 博士の名を冠して「コクラン共同計画」(The Cochrane Collaboration) と名づけられ、世界的に展開されている医療技術評価のプロジェクトである。

社会福祉・教育・刑事司法や学際領域としては、「キャンベル共同計画」(The Campbell Collaboration : C2)<sup>8)</sup>が

表1 エビデンス・グレーディング (治療の有効性に関して)<sup>1)</sup>

I	システマティック・レビューまたはメタ・アナリシス
II	1つ以上のランダム化比較試験による
III	非ランダム化比較試験による
IV	分析疫学的研究 (コホート研究や症例対照研究)
V	記述研究 (症例報告や症例集積)
VI	患者データに基づかない、専門委員会や専門家個人の意見

エビデンスの流れ (3局面)

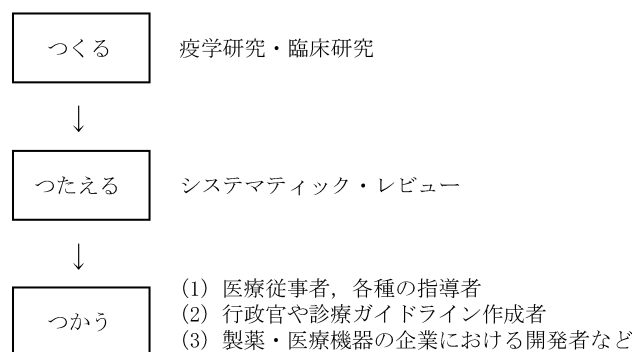


図1 エビデンスの流れとシステマティック・レビュー (津谷, 2003より作図)<sup>4)</sup>

よく知られている。2000年に発足し、コクラン共同計画の姉妹機関として密接な関係を有している。RCTとnRCTを明確に分け、SRによって、社会・教育施策や実務の効果に関する最善のエビデンスを知りたい市民、実務家、政策決定者、教員と学生・生徒、そして研究者に電子的に公表し、更新していく世界的な評価プロジェクトである。

#### 4. 論文の評価と質を高めるためのチェックリスト

エビデンスのグレードが高いRCTは、SRの重要な研究対象 (study population) である。しかし、論文における記述が曖昧だったり、不適切であれば、「非倫理的な医療行為を導く危険性があるとともに、科学的妥当性を欠くRCTは研究費や参加者を無駄に使うばかりでなく、後に続く研究を誤った方向へ導きうる」<sup>9)</sup>ことが指摘されている。世界でも数多くのRCTの論文が掲載される中で、こ

うした問題を回避するために各種のチェックリストが開発されている。

表2は、論文の質を高めるためのチェックリスト・声明と特徴の一覧である (略称はコンセンサスが得られたものではない)。最も著名なのが、22のチェック項目からなる「改訂版CONSORT声明」<sup>10)</sup>で、多くの学術雑誌から支持されている。世界的な5つの医学雑誌 (*Ann Inter Med*, *BMJ*, *JAMA*, *New Engl J Med*, *Lancet*) を含む「国際医学雑誌編集者委員会」(International Committee of Medical Journal Editors: ICMJE) では、査読前から記入漏れや不明確な記載がないようにするために、RCTの論文をウェブ上から投稿する際には、このチェックリストに基づいているかの確認シートがあり、それを完了しないと受け付けられない。MOHERら<sup>11)</sup>は、1996年に最初のCONSORT声明<sup>12)</sup>を採択した雑誌 (*BMJ*, *JAMA*, *Lancet*) とそうでない雑誌 (*NEJM*) に掲載された論文の質を声明発表の前後で比

表2 論文の質を高めるためのチェックリスト・声明と特徴

研究デザイン	名称または略称 (細目名)	項目数と特徴
システマティック・レビュー (SR: RCTによる)	QUOROM声明 <sup>13)14)</sup> (RCTのSR)	出版バイアスやメタ・アナリシスの方法、感度分析などSR特有の設問項目を含む18項目で構成されている。
ランダム化比較試験 (RCT)	改訂版CONSORT声明 <sup>9)10)</sup> (RCTの包括)	RCTの基本となるチェックリストであり、広くコンセンサスが得られている。22項目で構成されている。
	Cluster RCT <sup>15)</sup> * (多施設間のRCT)	倫理問題など、個人別の割付けの実施が難しい状況において、コミュニティーや病院、施設等の単位で行うRCTである。改訂版CONSORT声明の拡張版として22項目で構成されている。
	Noninferiority and equivalence RCT <sup>16)</sup> * (非劣勢・同等性)	第III相試験などで実施する非劣勢 (同等性) 試験は、プラセボとの比較ではなく、実薬と比較して劣らないことを実証する。こうしたデザインに対応した改訂版CONSORT声明の改変版である (22項目)。
	Harms in RCT <sup>17)</sup> * (有害事象)	有害事象を明記することで総合的にある介入の効果を評価することが主旨であり、改訂版CONSORT声明を改変した22項目から構成されている。
	CLEAR NPT <sup>18)</sup> (薬以外のRCT)	薬理学以外の研究ではブラインドが困難な場合が多く、介入者 (プロバイダー) のスキルによる影響が大きいなどの特異性がある研究も多い。そうした点を項目に加え、10項目と下位5項目からなるチェックリストである (表3)。
	STRICTA <sup>19)20)</sup> (鍼治療)	大項目が6で、臨床経験や刺鍼の詳細などの20項目から構成されている。鍼独自の項目が多い。
	PEDroスケール <sup>21)22)</sup> (理学療法)	PEDroの正式名称は、The Physiotherapy Evidence Databaseで、科学的根拠に基づいた理学療法を実施するための臨床研究文献のデータベースである。チェックリストは、同分野において広く用いられ、内的妥当性10、外的妥当性1の項目から構成されている。
	PEDroスケール改変版 <sup>23)</sup> (温泉療法)	原著の10項目に加えて、著者らが3項目を追加した温泉研究用のチェックリストである。
	Cochrane レビュー (転倒予防 <sup>24)</sup> ・漸増的筋力増強運動 <sup>25)</sup> ) と PEDroスケールの改変版 <sup>26)</sup> * (生活・運動指導)	転倒予防と筋力増強運動のコクラン・レビューで用いられた10項目とPEDroスケールの9項目、そして独自の2項目 (有害事象の有無、サンプル数) を組み合わせた15項目からなるチェックリストである。
非ランダム化比較試験 (nRCT)	TREND声明 <sup>27)</sup>	改訂版CONSORT声明に揃えた形式で22項目から構成されている。RCTは実施が困難なことが多く、現実的な介入研究ではnRCTsが実施しやすい。
システマティック・レビュー (SR: OSによる)	MOOSEチェックリスト <sup>28)29)</sup>	35項目で構成され、検索ストラテジーの細目では検索担当者の適格性 (図書館員や調査員) や入手可能な研究を収集する努力 (著者との連絡)、英語以外の論文の取り扱いなどが含まれている。
観察研究 (OS)	STROBE声明 <sup>30)31)</sup>	CONSORT声明の形式に揃え、コホート研究、症例対照研究、横断研究別に作成され、22項目から構成されている。

[注]

\* 筆者らが簡略したものであり、コンセンサスが得られた略称ではない。

記載・実施された  
項目数(40点満点)

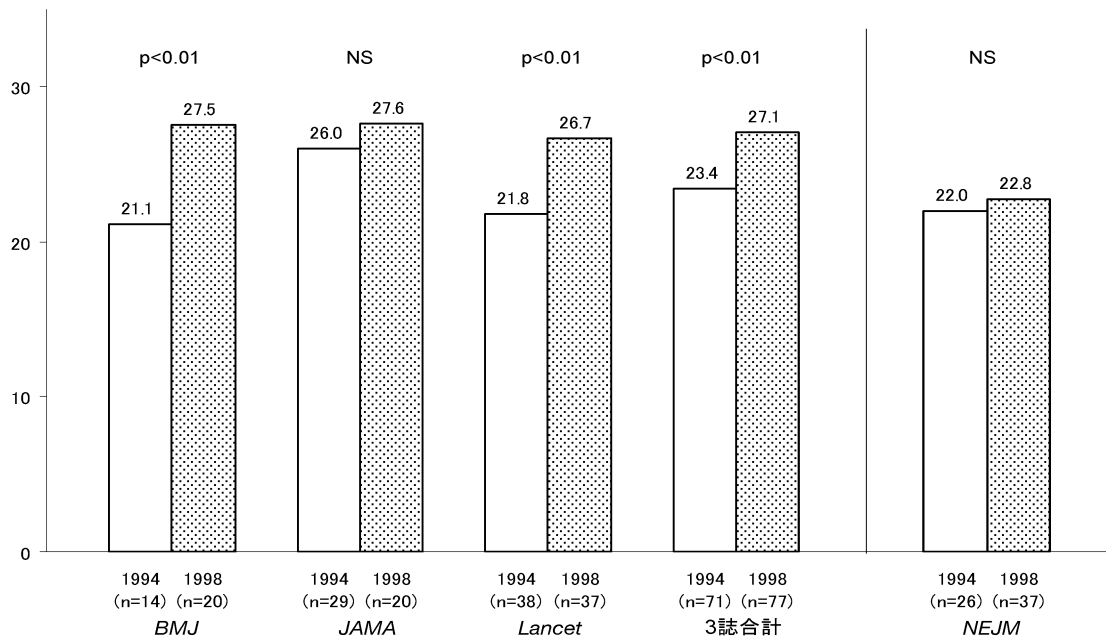


図2 CONSORT 声明の発表前後における論文の質の比較 (文献11の表から著者作図)

\* 1994年と1998年の前半期のランダム化比較試験が対象

\*\* BMJ, JAMA, Lancet は1996年からCONSORT 声明を採択

\*\*\* NEJM は1998年時点でCONSORT 声明を公式に採択していなかったため比較対照とした

較した(図2)。その結果、声明におけるチェック項目の記載数は、BMJとLancet及び3雑誌合計で有意に増加したが、NEJMでは変化が見られず、CONSORT 声明によって論文の質が高まったことを報告している。

表2のRCTにおいても、介入手法の違いによって合致しない項目があるのも事実で、それぞれの研究領域に応じて、変更していることが理解できる。nRCTとOSおよびSR(OS)は、RCTと比較すると実施しやすい研究デザインであり、実際の研究数も多い。

表3は、薬以外のRCTに対してのチェックリストである。薬理作用を調べる研究では、プラセボによってブラインド化が容易だったり、介入者のスキルの影響はほとんどない。しかし、外科療法、理学療法、運動療法などは、参加者(被験者)がその介入方法を身をもって知ってしまうため、ブラインド化が困難であったり、また介入者の経験やスキルによる影響が大きいことが問題となる。こうした研究に対応して開発されたのが、CLEAR NPTである。これらを参考にして、論文に漏れなく適切な情報を記載する必要がある。

## 5. 論文の質の吟味

チェックリストは、著者が論文を書くときに、質を高めることに寄与するとともに、編集者や査読者、SRの著者にとっては、論文の質評価に役立つ。

CHALMERSら<sup>32)</sup>は、1981年に研究の質を総合的に評価するために論文を得点化する方法を開発した。それ以後、各論文を得点評価する研究も数多くある一方で、無意味であ

るとする研究も多い。

HUWILER-MÜNTENERら<sup>33)</sup>は、論文全体としては良い報告であっても、方法の質(concealment: 隠蔽, ブラインド, intention-to-treat 分析)が必ずしも高いとは限らないことを60編のRCTの分析から報告している。つまり、合計得点ではなく、個々の項目の評価を行うことが重要であることを示唆している。

SOARESら<sup>34)</sup>は、悪性腫瘍での生存と患者のQOLを高めるための研究を推進するRadiation Therapy Oncology Groupによって行われたRCT 58編を分析した。その結果、事前のサンプルサイズの計算を実施していたのは44編(76%)だったが、論文中にそれを記載していたのは、わずか9編(14%)であり、同様に隠蔽については58編すべてで実施されていたが、記載してあったのは24編(41%)だったことを報告している。

論文には、必ずしも実施したことをすべて記載しているとは言えず、臨床研究登録の研究計画書を調べたり、研究者に直接問い合わせることも重要であるとし、合わせてチェックリストに基づいて記載すべきことを強調している。

則内ら<sup>35)</sup>は、特定保健用食品の有効性を検証したRCT 36編の質を評価した結果、症例数の決定方法やランダム化については、記載のないもののがかなり多く、全体的にも32項目中、平均15.2で低い水準にあった(表4)。この報告も、RCTの質を担保するために改訂版CONSORT 声明に準拠すべきことを強調している。

### 6. 臨床研究登録の必要性

パブリケーション・バイアスは、エビデンスに大きな悪影響をもたらすひとつである。典型的な例としては、ある介入方法に対して「効果あり」といった正の結果の論文は受理されやすく、負の結果のものは採用されにくい。そのため、正の結果ばかりが強調されてしまい、エビデンスの誤認が生じうるとのことである（図3）。

1990年以前では、雑誌に掲載された研究結果は、関係研究者やシステマティック・レビューの分析の対象となるが、そうではない研究、例えば、研究は実施したが仮説どおりではなく、論文としていない研究や学術誌に投稿したが不採用のままの論文の結果が表面に出てこない状況にあった。1999年に世界に先駆けて、*Lancet*<sup>36)</sup>などの医学雑誌は、事前の研究登録を推奨した。さらに、2004年には、ICMJEが、臨床試験の登録（clinical trial registry：CTR）を論文投稿の条件とした<sup>37)</sup>。またWHOも2005年にジュネーブ会議を開催し、登録をポリシーとし基準を定めた<sup>38)</sup>。

日本においても、公表バイアスを回避するために、「臨床試験登録」(UMIN-CTR)<sup>39)</sup>が開始され、年度ごとに登録件数も増加している。こうした臨床試験の事前登録は、パブリケーション・バイアスの減少、研究倫理、臨床試験参加の推進の3つが目的であるが、SRのレビュアーが、包括的・網羅的に漏れなく論文（研究）を収集する助けにもなる。

SRにおける情報バイアスとして、レビュー・バイアスがある。VERHAGENら<sup>40)</sup>は、よりインパクトファクターの高い雑誌や著名な雑誌に掲載された論文の質は、信頼できる査読をパスしているため、高く評価してしまう可能性があることを指摘している。図3に示したように、論文として掲載されていなくても、実施した研究結果をすべて包括してのSRが望まれ、CTRがそれをカバーしてくれるのである。

各種の研究領域においても、人を対象とした介入研究で、その介入方法や評価に侵襲性を伴うような場合にはUMIN-CTRへの登録が必要である。

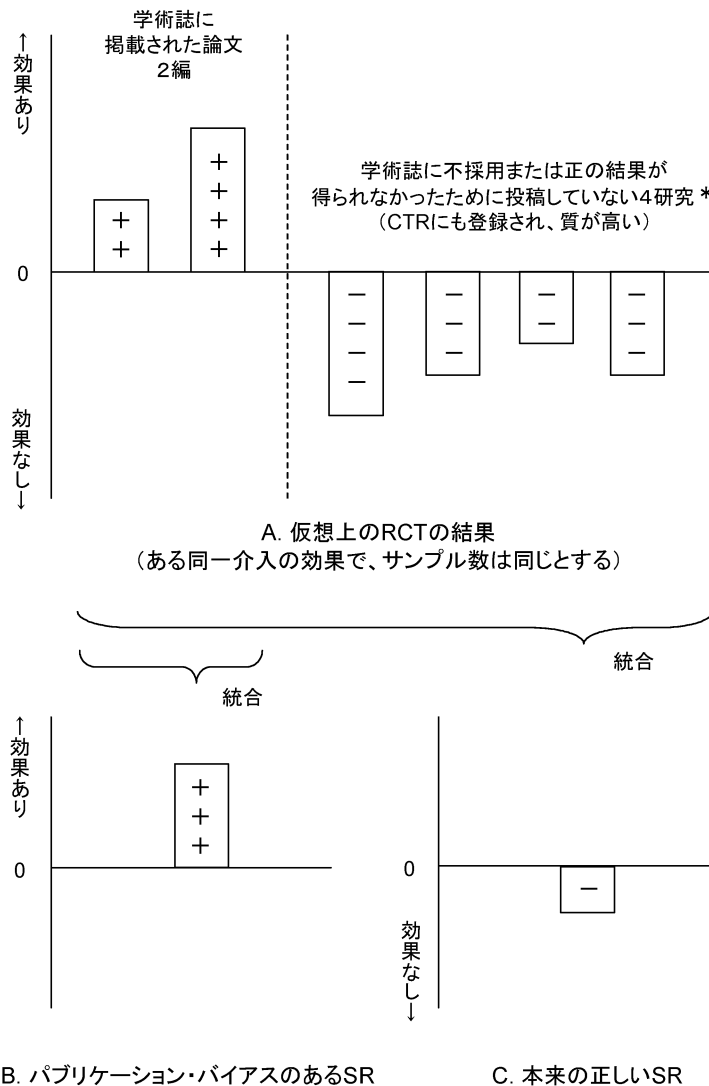


図3 パブリケーション・バイアスの典型例

\* 現在では正の結果が得られなかった論文でも質が高ければ学術誌に掲載される傾向にある。

表 3 非薬理学研究におけるランダム化比較試験の質評価チェックリスト (文献 18 を著者らが翻訳・作表)

項目	記載状況
1. 割付順番の作成は適切だったか?	はい; いいえ; 不明確
2. 治療の割付は隠蔽されたか?	はい; いいえ; 不明確
3. 各群に施した介入の詳細は利用可能か? <sup>a</sup>	はい; いいえ; 不明確
4. 各部門における介入者の経験と技能 <sup>b</sup> は適切であったか? <sup>c</sup>	はい; いいえ; 不明確
5. 参加者(すなわち患者)のアドヒレンスは量的に評価されたか? <sup>d</sup>	はい; いいえ; 不明確
6. 参加者は十分に盲検化されたか?	はい; 盲検化不可能; 可能だがしていない; 不明確
6.1. 参加者が十分に盲検化されていない場合	
6.1.1. 他のすべての治療(すなわち介入に関わるもの)はどのランダム化された各群でも同様であったか?	はい; いいえ; 不明確
6.1.2. ドロップアウトはランダム化された各群で同様であったか?	はい; いいえ; 不明確
7. ケアプロバイダーや参加者を世話する人は十分に盲検化されたか?	はい; 盲検化不可能; 可能だがしていない; 不明確
7.1. ケアプロバイダーが十分に盲検化されていない場合	
7.1.1. 他のすべての治療(すなわち介入に関わるもの)はどのランダム化された各群でも同様であったか?	はい; いいえ; 不明確
7.1.2. ドロップアウトはランダム化された各群で同様であったか?	はい; いいえ; 不明確
8. アウトカム評価者は、主要なアウトカムの評価に際して十分に盲検化されたか?	はい; 盲検化不可能; 可能だがしていない; 不明確
8.1. アウトカム評価者が十分に盲検化されていない場合、確認バイアス(系統誤差)を避けるために特別な方法を用いたか? <sup>e</sup>	はい; いいえ; 不明確
9. フォローアップ・スケジュールは各群で同様だったか? <sup>f</sup>	はい; いいえ; 不明確
10. 主要アウトカムはITT分析理論に基づいて解析されたか?	はい; いいえ; 不明確

<sup>a</sup>これらのデータが記載されたり、各部門(先行研究の参照、オンライン付録など)で利用可能な場合には、この回答は「はい」とすべきである。

<sup>b</sup>ケアプロバイダーの経験や技能は治療者依存の介入についてのみ評価される。(治療の成功は直接的にプロバイダーの技能と結びつけられる)他の治療に関しては、この項目は適切ではなく、チェックリストから除外されるか、「不明確」と回答されるべきである。

<sup>c</sup>適切な経験・技能は公表されたデータ、先行研究、ガイドライン、準備期間、または専門家集団にしたがって決定されるべきである。そして、調査開始前に各研究部門のプロトコールの中で明確にされるべきである。

<sup>d</sup>治療の順守度は、反復介入(例、手術のような単発の治療とは対照的ないくつかのセッションからなる理学療法など)を必要とする治療についてのみ評価される。単発の治療については、この項目は適切ではなく、チェックリストから除外されるか、「不明確」と回答されるべきである。

<sup>e</sup>主要アウトカムが客観的ないし頑強な場合、アウトカムが盲検化された、ないしは少なくとも独立したエンドポイント管理者によって評価された場合、標準化された方法で評価するように訓練された独立した評価者によって行われた場合、アウトカム評価者が研究目的・仮説に対して盲検化された場合には、この項目に対する回答は「はい」であるべきである。

<sup>f</sup>この項目は、フォローアップが論点の一部となる試験では適切ではない。

(例: ガンの再発に対するフォローアップの頻度を評価する試験)

このような状況では、チェックリストから除外されるか、「不明確」と回答されるべきである。

附記: 本研究の実施にあたり、平成 19 年度厚生労働科学研究費補助金(循環器疾患等生活習慣病対策総合研究事業)、「温泉利用と生活・運動・食事指導を組み合わせた職種別の健康支援プログラムの有効性に関する研究(主任研究者: 上岡洋晴)」の一部として実施した。

文献の検索と収集に際しては、司書兼サーチャーとしてご協力いただきました東邦大学習志野メディアセンターの眞喜志まり氏と東京厚生年金病院図書室の山田有希子氏に感謝いたします。

表 4 CONSORT 声明に基づく「特定保健用食品」の有効性検証の質評価（文献 35 を著者らが一部改変作表）

見出し(項目数)	項目	合計* n=36		
タイトルと抄録(1項目)	参加者の介入群への配置	7	(19%)	
はじめに(1項目)背景	科学的背景と合理的背景	36	(100%)	
方法(18項目)	参加者	参加者の適格条件	32 (97%)	
		データが収集された状況と場所	7 (19%)	
		各群に意図された介入の詳細	36 (100%)	
	介入	実際にいつ実施されたか	5 (14%)	
		実際にどのように実施されたか	36 (100%)	
		特定の目的と仮説	34 (94%)	
	目的	主要・副次的アウトカム評価項目の定義	36 (100%)	
		測定の質を向上させる方法	19 (53%)	
		目標症例数の設定方法	0 (0%)	
	症例数	中間解析と中止基準の説明	0 (0%)	
		割付け順番を作成した方法	0 (0%)	
		ランダム化	割付けの制限の詳細	5 (14%)
			割付けの実施方法と隠蔽	1 (3%)
			割付けの実施者	1 (3%)
		ブラインディング・マスキング	割付け状況のブラインド化	25 (69%)
			ブラインド化の成功の評価	2 (6%)
		統計学的手法	群間比較に用いられた統計学的手法	35 (97%)
			追加的解析の手法	8 (22%)
結果(9項目)		参加者の流れ	各段階を通じた参加者の流れ	17 (47%)
			プロトコル逸脱の人数と理由	19 (53%)
		募集	募集期間と追跡期間	36 (100%)
	ベースラインのデータ	各群のベースラインのデータ	30 (83%)	
	解析された人数	各解析の各群の参加者数(分母)	28 (78%)	
		ITT解析かどうか	2 (6%)	
	アウトカムと推定	介入のエフェクトサイズと精度(95%信頼区間)	2 (6%)	
		補助的解析	サブグループ解析や調整解析	4 (11%)
	有害事象	各群の重要な有害事象	30 (83%)	
	考察(3項目)	解釈	結果の解釈(研究の限界)	21 (58%)
一般化可能性		結果の一般化可能性	1 (3%)	
全体としてのエビデンス		結果の包括的解釈	33 (92%)	
全体(32項目)		15. 2項目		

以下、特定保健用食品の有効性を検証①—③までのRCTを評価している。

①2003年に許可された製品について(12報)

②「お腹の調子を整える」等の表示をした食品中の食品乳酸菌類を含む食品について(5報)

③「健康・栄養食品研究」誌上に論文が報告された製品について(23報)

\*合計については重複している論文を除いてある

#### 引用文献

- 財団法人厚生統計協会：図解 国民衛生の動向，p 12，2000.
- 丹後俊郎・上坂浩之編，2006. 臨床試験ハンドブック，朝倉書店，東京，3-5.
- LAST, J.M. 編，日本疫学会誌，2000. 疫学辞典第3版，財団法人日本公衆衛生協会，東京，183.
- 津谷喜一郎 (2003) EBM におけるエビデンスの吟味. *Therapeutic Research*, 24, 1415-1422.
- 津谷喜一郎 (2003) エビデンスを調べる—systematic review の現状—. *臨床薬理*, 34 (4), 210-216.
- 丹後俊郎 (2002) メタ・アナリシス入門，朝倉書店，東京，36-42.
- コクラン共同計画：アクセス，<http://www.cochrane.org/>
- キャンベル共同計画：アクセス，<http://www.campbell.gse.upenn.edu/>
- 津谷喜一郎，小島千枝，中山健夫訳・解説 (2002) CONSORT 声明：ランダム化並行群間比較試験報告の質向上のための改訂版勧告. *JAMA* 〈日本語版〉 118-124.
- MOHER, D., SCHULZ, K.F., ALTMAN, D.G., *et al.*, (2001) The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 285, 1987-1991.
- MOHER, D., JONES, A., LEPAGE, I. *et al.*, (2001) Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA* 285, 1992-1995.
- BEGG, C., CHO, M., EASTWOOD, S., *et al.*, (1996) Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 276, 637-639.
- MOHER, D., COOK, D.J., EASTWOOD, S., *et al.*, (1999) Improving the quality of reports of meta-analyses of randomized controlled trials: the QUOROM statement. *Lancet* 354, 1896-1900.
- 津谷喜一郎，折笠秀樹，熊井智子訳，[http://homepage3.nifty.com/cont/QUOROM\\_Statement/QUOROM\\_chk\\_Jap.pdf](http://homepage3.nifty.com/cont/QUOROM_Statement/QUOROM_chk_Jap.pdf)
- CAMPBELL, M.K., ELBOURNE, D.R., ALTMAN, D.G., *et al.*, (2005) CONSORT statement: extension to cluster randomized trials. *BMJ* 328, 702-708.
- PIAGGIO, G., ELBOURNE, D.R., ALTMAN, D.G., *et al.*, (2006) Reporting of noninferiority and equivalence randomized trials: An extension of the CONSORT Statement. *JAMA* 295, 1152-1160.
- LOANNIDIS, J.P.A., EBVANS, S.J.W., GOTZSCHE, P.C., *et al.*, (2004) Better reporting of harms in randomized trials:

- An extension of the CONSORT Statement. *Ann Intern Med* 141, 781-788.
- 18) BOUTRON, I., MOHER, D., TUGWELL, P., *et al.*, (2005) A checklist to evaluate a report of a nonpharmacological trial (CLEAR NPT) was developed using consensus. *J Clin Epidemiol* 58, 1233-1240.
  - 19) MACPHERSON, H., WHITE, A., CUMMINGS, M., *et al.*, (2001) Standards for reporting interventions in controlled trials of acupuncture: the STRICTA recommendations. *Complement Ther Med* 9, 246-249.
  - 20) 津嘉山洋, 山下 仁 (2002) 鍼の臨床試験におけるデザインと報告に関する統一規格: STRICTA グループとIARFの勧告. *全日鍼灸会誌*, 52, 582-586.
  - 21) Physiotherapy Evidence Database: Access, <http://www.pedro.fhs.usyd.edu.au/index.html>
  - 22) 下坂 充 (2001) EBPTにおけるデータベースの活用法. *理学療法*, 18, 1036-1042.
  - 23) 上岡洋晴, 黒柳律雄, 小松泰喜, 上内哲男, 高橋美絵 (2006) 温泉の治療と健康増進の効果に関する無作為化比較試験のシステマティック・レビュー. *日温気物医誌*, 69, 155-166.
  - 24) GILLERSPIE, L.D., GILLESPIE, W.J., *et al.*, (2004) Interventions for preventing falls in elderly people (Cochrane Review). The Cochrane Library, Issue 3, Chichester, UK: John Wiley & Sons, Ltd.
  - 25) LATHAM, N., ANDERSON, C., BENNETT, D., *et al.*, (2004) Progressive resistance strength training for physical disability in older people (Cochrane Review). The Cochrane Library, Issue 3, Chichester, UK: John Wiley & Sons, Ltd.
  - 26) 高橋美絵, 上岡洋晴, 津谷喜一郎ら (2007) 中高年者の健康増進を目的としたランダム化比較試験による運動・生活指導介入のシステマティック・レビュー: 介入研究の課題と介入モデルの検討. *日本老年医学会誌*, 44, 403-414.
  - 27) JARLAIS, D.C., LYLES, C., CREPAZ, N., *et al.*, (2004) Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions. *Am J Public Health* 94, 361-366.
  - 28) STROUP, D.F., BERLIN, J.A., MORTON, S.C. *et al.*, (2000) Meta-analysis of observational studies in epidemiology. *JAMA* 283, 2008-2012.
  - 29) 中山健夫, 津谷喜一郎訳・解説 (2002) 疫学分野における観察研究のメタ・アナリシス報告のための提案. *JAMA* 〈日本語版〉 52-58.
  - 30) ELM, E., ALTMAN, D.G., EGGER, M., *et al.*, (2007) The strengthening the reporting of observational studies in epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Ann Intern Med*, 147, 573-577.
  - 31) 上岡洋晴, 津谷喜一郎 (訳) (2008) 疫学研究における観察研究の報告の強化 (STROBE 声明): 観察研究の報告に関するガイドライン. 臨床研究のための国際的ルール (中山健夫, 津谷喜一郎監訳), ライフサイエンス出版, 東京 (印刷中).
  - 32) CHALMERS, T.C., SMITH, H.J., BLACKBURN, B. *et al.*, (1981) A method for assessing the quality of a randomized control trial. *Control Clin Trials* 2, 31-49.
  - 33) HUWILER-MÜNTENER, K., JINI, P., JUNKER, C., *et al.*, (2002) Quality of reporting of randomized trials as a measure of methodologic quality. *JAMA* 287, 2801-2804.
  - 34) SOARES, H.P., DANIELS, S., KUMAR, A., *et al.*, (2004) Bad reporting does not mean bad methods for randomized trials: observational study of randomized controlled trials performed by the Radiation Therapy Oncology Group. *BMJ* 328, 22-24.
  - 35) 則内友博, 大橋靖雄 (2006) 特定保健用食品の有効性検証試験の論文報告—CONSORT 声明を用いた質評価. *薬理と治療*, 34, 1039-1051.
  - 36) HORTON, R. and SMITH, R. (1999) Time to register randomized trials. The case is now unanswerable. *Lancet* 354, 1138-1139.
  - 37) 第25回日本臨床薬理学会年会シンポジウム12 [5人のスピーカーによる講演記録集] (2005) 臨床試験の登録と結果の公開 (ポジティブ, ネガティブを含めて). *臨床医薬*, 21, 27-62.
  - 38) 津谷喜一郎 (2005) 臨床試験登録に関する「オタワ声明」と「ジュネーブ会議」の動向. *薬理と治療*, 33, 543-566.
  - 39) 医学情報大学病院医療情報ネットワーク, <http://www.umin.ac.jp/ctrindex-j.htm>
  - 40) VERHAGEN, A.P., HENRICA, C.W.V., ROBERT, A.B., *et al.*, (2001) The art of quality assessment of RCTs included in systematic reviews. *J Clin Epidemiol* 54, 651-654.



# The Current International View of the Enhancement of the Quality of Scientific Publications and Improvement of the Design of Clinical and Epidemiological Research in Humans

By

Hiroharu KAMIOKA\*, Kiichiro TSUTANI\*\*, Yukari KAWANO\*\*\*,  
Yoshiteru MUTOH\*\*\*\*, Nobuyoshi SHIOZAWA\*\*\*\*\*,  
Yoshihisa MIYAMOTO\*\*\*\* and Takuya HONDA\*\*\*\*

(Received October 5, 2007/Accepted January 18, 2008)

**Summary** : This narrative reviews the means of effectively evaluating evidence in the outcome of clinical trial and quality of health care. The grading of evidence in clinical and epidemiologic studies has been provided to improve the study design and to raise the quality of publications. Check lists and guidelines are developed to reduce publication bias, and clinical trial registry has the potential to facilitate ethics and to increase participation of clinical trials. The International Committee of Medical Journal Editors has proposed that only clinical trials are published that are registered at inception. These trends are to be appreciated when epidemiologic or clinical study is conducted.

**Key words** : evidence grading, epidemiology, clinical trials, study design, checklists, statement

---

\* Laboratory of Physical and Health Education, Faculty of Regional Environment Science, Tokyo University of Agriculture

\*\* Department of Drug Policy and Management (DPM), Graduate School of Pharmaceutical Sciences, The University of Tokyo

\*\*\* Department of Nutritional Sciences, Faculty of Applied Bio-Science, Tokyo University of Agriculture

\*\*\*\* Department of Physical and Health Education, Graduate School of Education, The University of Tokyo

\*\*\*\*\* Department of Longevity and Social Medicine, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences