

<総説>

エビデンスの構築と研究方法論の向上を目的とした論文の質評価に関する考察  
— 学会誌「レジャー・レクリエーション研究」における  
1993-2007年までの疫学的論文を対象として —

上岡洋晴<sup>1</sup> 鈴木英悟<sup>2</sup> 栗田和弥<sup>3</sup> 本多卓也<sup>4</sup>

**A narrative review on quality estimation of studies for  
evidence construction and improvement of research methodology:  
epidemiological studies published in a scientific journal  
"Journal of Leisure Recreation Studies" between 1993 and 2007**

Hiroharu Kamioka<sup>1</sup>, Eigo Suzuki<sup>2</sup>, Kazuya Kurita<sup>3</sup>, Takuya Honda<sup>4</sup>

**Abstract**

In this study, we evaluated the quality of articles on human subjects published in a scientific journal, "Journal of Leisure Recreation Studies", using various kinds of checklists in order to identify possible problems with establishing the evidence.

We collected and studied articles published in "Journal of Leisure Recreation Studies" in the 15 years between 1993 and 2007. The criteria for adopting articles was their hypothesis testing strategies, with study designs including interventional studies using randomized controlled trials (RCT) and non-randomized controlled trials (nRCT), and observational studies such as cross-sectional studies (Cr), case-control studies (Ca) and cohort studies (Co). The interventional studies also included those performed without relevant control groups. Quality of the articles was evaluated using the two checklists: "TREND Statement Checklist" comprised of 58 items, and "STROBE Statement Checklist" comprised of 38 items.

Five of the 42 articles published in these 15 years were screened out, comprising 11.9% of the total. The following major solutions were proposed in order to improve the quality of the articles: 1) search for earlier domestic and international earlier studies utilizing various databases, 2) defining leisure activities and recreation, 3) ethical consideration in research activities, and 4) citation of articles that complied with the statement checklists whose international consensus has been established.

- 
- 1 東京農業大学地域環境科学部身体教育学研究室  
Laboratory of Physical and Health Education, Faculty of Regional Environment Science, Tokyo University of Agriculture
  - 2 東海大学体育学部  
Faculty of Physical Education, Tokai University
  - 3 東京農業大学地域環境科学部造園科学科  
Department of Landscape Architecture, Faculty of Regional Environment Science, Tokyo University of Agriculture
  - 4 東京大学大学院教育学研究科身体教育学講座  
Department of Physical and Health Education, Graduate School of Education, The University of Tokyo

## 1. 緒言

1990年代後半から「科学的根拠に基づいた医療 (evidence based medicine)」や「科学的根拠に基づいた健康政策 (evidence based health policy)」など、「科学的根拠に基づいた…」という用語が頻繁に用いられるようになった。このエビデンス<sup>2)</sup>を構築する世界的な流れは、医学分野に留まらず、保健・福祉・教育・刑事司法など、人を対象としてその効果を明らかにしようとする研究分野全体 (広義の疫学的研究) に伝わってきている<sup>1)</sup>。そして、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 (格付け)<sup>2)</sup>がなされている (表1)。

このグレーディングは、ヒトゲノムなどの基礎医学ではなく、実際の人間を対象とした疫学研究や臨床研究などの研究デザインが対象となる。レジャー・レクリエーションは、人文学・社会学・自然環境学・芸術学などを包括する極めて広い分野・領域であるが、主として、レジャー活動やレクリエーションが、心身に及ぼす効果や教育効果、あるいは治療効果を探るような研究の場合には、この考え方が合致すると考えられる。エビデン

ス・グレーディングの理解は、研究者自身が予め計画段階において、明らかにしようとする仮説を立証するのに役立つ、現実的な実施可能性と照らし合わせて、方法論を選択する助けとなる (表1の具体例を参照)。

「レジャー活動」と「レクリエーション」に関するランダム化比較試験 (Randomized Controlled Trial: RCT) の2006年のシステマティック・レビュー (Systematic Review: SR) では、1990年から1996年までの期間に世界で掲載されたRCT論文は、わずかに3編だったことが報告されている<sup>3)</sup>。

しかし、RCTは少ないものの、非ランダム化比較試験 (nRCT) や横断研究 (Cross-Sectional Study: Cr)、症例対照研究 (Case-Control Study: Ca)、コホート研究 (Cohort Study: Co) などの観察研究 (Observational Study: Os) は行われていると考えられる。RCTは、研究者が仮説として準備したプログラムを実施する群と実施しない群にランダムに割付して、その効果を明らかにしようとする方法である。nRCTは、ランダムではなく、介入群を希望する参加者は、そちらにして、希望しない群をコントロール群にするなど、最初から参加

表1 エビデンス・グレーディング (引用文献3から転載)

- 
- |     |                             |
|-----|-----------------------------|
| I   | システマティック・レビュー (メタ・アナリシスを含む) |
| II  | 1つ以上のランダム化比較試験による研究         |
| III | 非ランダム化比較試験による研究             |
| IV  | 分析的疫学研究 (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  |
- 

[注] レジャー活動やレクリエーションの研究に当てはめた例 (架空)

- |     |                                                                                                                                                                                                   |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| I   | IIに基づく複数の研究結果を網羅的に収集し、メタ・アナリシスという統計手法に基づき統合するとともに、批判的吟味を加えて、解釈や一般化可能性 (外的妥当性)、全体的なエビデンスを示すこと。                                                                                                     |
| II  | あるレクリエーションをさせる群と何もさせない群にランダムに割付し、その効果を見ること。                                                                                                                                                       |
| III | あるレクリエーションをさせる群と何もさせない群に研究者の意図に基づいて割付し、その効果を見ること。                                                                                                                                                 |
| IV  | 1) ある市の全小学校において、ボーイスカウトに入っている子どもとそうでない子どもに分けて10年間追跡し、10年後時点でボランティア活動を行っている者の比率を比較すること。(コホート研究)<br>2) ある小学校において、アウトドア活動を1年に2日以上行っている児童 (実践群) と、1年間に1日以下しか行っていない児童 (非実践群) に分けて、体力テストの結果を比較する。(横断研究) |
| V   | 数例 (統計解析ができない程度) のレジャー活動やレクリエーションの報告や、実施前後の客観的データの比較、参加者の主観的な態度や心の変化などの記述。(事例研究)                                                                                                                  |
| VI  | 研究データに基づかない、専門委員会や専門家個人の意見                                                                                                                                                                        |

者や研究者の意図を反映した割付がなされてのデザインである。Crは、1回の調査において多数のサンプルを集め、そのデータを仮説に基づいて複数の群に分けて、その差異を明らかにするようなデザインである。Caは、患者と非患者などで複数の群を設定し、仮説としている変数について群間比較をするようなデザインである。Coは、大集団を複数年にわたり追跡するようなデザインである。

前述のすべての研究デザインにおいては、国際的にコンセンサスが得られている論文の質を高めるための声明やチェックリストが開発されている。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>4) 5)</sup>で、多くの学術雑誌から支持されている。とくに世界的な医学雑誌 (Ann Inter Med, BMJ, JAMA, New Engl J Med, Lancet) 11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、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。つまり、このチェックリストに基づいて予め研究計画・実施・分析・執筆をすれば、結果として質が高まることを意味している。こうしたチェックリストに基づいて論文の評価を行い、レジャー・レクリエーション分野における研究のウィーク・ポイントや論文執筆の課題となる諸点を明らかにして、より良い方法論の確立を目指すことは重要である。

ところで、「レジャー・レクリエーション研究」のように国レベルの学会誌は、当該研究分野の学術成果の最高の集積の場であり、学会員にとって最新の研究を学ぶテキストでもある。また、学会関係者以外の研究者にも読まれることから、それを意識した論文の設えも必要である。それは、人を対象とした研究の成果を自然科学の共通言語として、社会医学や臨床医学などの分野の研究者にも理解できる方法で示すことである。このことが、レジャー・レクリエーションの優位性や重要性を正しく理解してもらうことに繋がるのである。

また、レジャー・レクリエーションの学問の大部分は、人の幸福(教育効果、健康増進効果、治

療効果など)に直接的に寄与することを目的としているので、方法論の整理がより重要だと考えられる。ところが、「レジャー・レクリエーション研究」において、人を対象として教育や健康増進効果、治療効果の方法論を議論した報告は、過去15年間にほとんど見られない。

そこで、本研究は、学術雑誌「レジャー・レクリエーション研究」で、1993年から2007年に掲載された人を対象とした論文を網羅的に収集して、各種のチェックリストを用いて論文の質評価を行うとともに、エビデンス構築のための課題点の整理をすることを目的とした。

## 2. 方法

### (1) 論文収集

論文の収集として、日本学術会議登録雑誌として再登録<sup>6)</sup>された1993年から、直近の2007年までの15年間に発行された学会雑誌「レジャー・レクリエーション研究」(以下、レジャー・レク誌)を対象論文とした。論文採用の適格基準と除外基準を表2に示した。仮説実証型の研究デザインとしては、RCT、nRCTの介入研究、観察型では、Cr、Ca、CoのOsとした。ただし、介入研究では、対照群を用いていない介入研究も含めた。人を研究対象としているが、運動生理学に類する実験的な研究、学会抄録(発表論文集)や学会報告記、特集などは除外した。

### (2) エビデンス・テーブル

介入研究では、世界的に標準の構造化抄録8項目からなる「目的、研究デザイン、セッティング、対象者、介入内容、主なアウトカム評価項目、主な結果、結論」で示した。Osは、「目的、研究デザイン、セッティング、対象者、主なアウトカム評価項目、主な結果、結論」の7項目で示した。

### (3) 研究の質評価

RCTでは「改訂版 CONSORT 声明チェックリスト」<sup>4) 5)</sup>、nRCTでは「TREND 声明チェックリスト」<sup>7) 8)</sup>、Cr、CaおよびCoでは「STROBE 声明チェックリスト」<sup>9) 10)</sup>を用いた。それぞれ、大項目22から構成されているが、項目内の詳細項目(下位項目)も分割して独立項目として評価した。TREND 声明チェックリストは58項目、STROBE 声明チェックリストは32項目であった。



林活動が自閉症者にどのような影響をもたらすかを調べた研究で、治療方法として効果のある可能性を報告している。表7は、活動前の疲労度別に見たスポーツ活動の効果についての研究で、事前に疲労感が皆無でスポーツ活動を行うと疲労度は低く維持されたままで、レクリエーションスコア

(満足度)も高くなることを報告している。表8は、大学生のレジャーにおける退屈感を調べた研究で、レジャー能力、集団の状況、性格、孤独感、生活における時間意識が退屈感と関連があることを報告している。

表4 エビデンス・テーブル(1):引用文献14

1. 研究題目  
環境教育プログラムを導入したキャンプの効果：参加者の自然に対する態度、イメージに着目して
2. 代表著者名(出版年)  
岡村泰斗(1996)
3. 目的  
環境教育プログラムを含んだキャンプ参加者(実験群[注])の自然に対する態度、イメージのキャンプ中の変化を、同プログラムを含まないキャンプ参加者(対照群)と比較して明らかにする。
4. 研究デザイン  
非ランダム化比較試験
5. セッティング  
実験群：1994年7月27日から8月9日の期間(13泊14日) 国立那須甲子少年自然の家  
対照群：1994年7月31日から8月10日の期間(10泊11日) 静岡県立朝霧野外活動センター
6. 対象者  
実験群：国立那須甲子少年自然の家主催の那須甲子アドベンチャーキャンプに参加した小中学生46名(男子56.5%、女子43.5%：小学5年生11名、小学6年生22名、中学1年生10名、中学2年生3名)  
対照群：静岡県教育委員会主催の静岡県フロンティアアドベンチャーキャンプに参加した小中学生80名(男子57.5%、女子42.5%：小学5年生23名、小学6年生38名、中学1年生19名)
7. 介入内容  
実験群：仲間作りハイキング、サバイバル生活、環境教育プログラム\*、縦走登山で、少年自然の家の専門教員、野外運動専攻の大学院生などが指導した。  
対象群：冒険オリエンテーリング、班別活動、個人別選択プログラム、サバイバルハイクで、野外運動を専門とする大学教官、静岡県キャンプカウンセラーが指導した。  
\*ブナ林をテーマに学習場面として、「植物、動物、水、土」の4つの自然資源を利用したプログラムを設定し、ブナ林に関するビデオ鑑賞、実践活動、集団露営、発表を4日間実施した。
8. 主なアウトカム評価項目  
自然に対する態度：「植物、動物、水、土」についての35項目で7段階のリッカートスケールを用いた。  
自然に対するイメージテスト：刺激語を「植物、動物、水、土」とした10の形容詞対とし、各5段階のリッカートスケールを用いた。
9. 主な結果  
態度得点において、実験群は植物、水、土で介入後に有意に向上し、1ヵ月後まで維持され、対照群とも有意な差がみられた。イメージ得点においても、実験群では植物、水、土で有意に向上し、1ヵ月後まで維持され、対照群とも有意な差がみられた。
10. 結論  
環境教育プログラムは、自然に対する望ましい態度変容、イメージ変容に効果的であった。

[注] 実験群を介入群、対照群をコントロール群と読み替える。

表5 エビデンス・テーブル(2): 引用文献15

1. 研究題目  
キャンプカウンセラーの性役割がキャンパーの性役割意識に及ぼす影響
2. 代表著者名(出版年)  
関智子(1996)
3. 目的  
1) カウンセラーに対するキャンパーの性役割意識構造を明らかにする。2) カウンセラーに対するキャンパーの性役割意識がキャンプ経験によってどのように変化するかを明らかにする。
4. 研究デザイン  
対照群のない介入試験
5. セッティング  
1993年7月11日から12日の事前研修および同年8月1日から19日の期間に行われた静岡県フロンティア・アドベンチャー・キャンプ
6. 対象者  
小中学生76名(男子42名、女子34名)とカウンセラー8名(男性4名、女性4名)
7. 介入内容  
キャンパーの学年・男女をほぼ均等に配分し、1班8-9名の8班編成とし、それぞれカウンセラーが1名ずつ割り当てた。主なキャンププログラムは、冒険ウォークラリー、1泊2日のビバーク(サバイバルハイク)と、自由に2種目選べるお好み活動(パラグライダー、川遊び、マウンテンバイク、シャワークライミング、ロッククライミング、草木染め、バターづくり)などであった。キャンプ全体を通じて、カウンセラーおよびキャンパーの性役割上の操作は行わなかった。しかし、お好み活動では、マウンテンバイク、川遊び、パラグライダー、シャワークライミング、ロッククライミングなどの動的活動は男性スタッフと少数の女性スタッフがを行い、草木染めやバター作りなどの静的な活動は主として女性スタッフが行った。
8. 主なアウトカム評価項目  
カウンセラーの条件・資質として必要な49項目について、男女どちらの方が当てはまるかの質問紙。「男子である:-2点」、「どちらかといえば男子である:-1点」、「どちらともいえない:0点」、「どちらかといえば女子である:1点」、「女子である:2点」とした。
9. 主な結果  
男性役割に属する「活動性・行動力」の性役割意識は、担当カウンセラーの性別に関わらず男女共性役割観へと変化した。女性役割に属する「気づかい・細やかさ」に対する女性カウンセラー班の女子キャンパーの性役割意識は男女共性役割へと変化した。
10. 結論  
カウンセラーは、キャンパーの性別に付随する固定的な役割観を理解するとともに、カウンセラー自身の性役割が及ぼす影響力を考慮して指導を行うべきである。

表6 エビデンス・テーブル(3): 引用文献16

1. 研究題目  
The possibility of forest activities in the autistic disabilities treatment by utilizing the rural forest  
(日本語名: 自閉症療育における里山を利用した山林活動の可能性)
2. 代表著者名(出版年)  
Uehara Iwao(1999)
3. 目的  
森林作業と散策を中心としたレクリエーションからなる森林活動が自閉症者に及ぼす影響を明らかにする。
4. 研究デザイン

表6：続き

- 対照群のない介入試験
5. セッティング  
施設の周囲の落葉樹の里山で、1995年5月から1997年8月の期間に行われた。
  6. 対象者  
長野県北安曇野郡池田町の自閉症療育施設「白樺の家」に入所し、重篤な身体障害のない22名（男性17名、女性5名、開始時の平均年齢19.5歳、標準偏差3.6歳）
  7. 介入内容  
森林作業（丸太への椎茸の菌の挿入）とレクリエーション（主にハイキング）の比率は3:2から4:1で年間を通じて月曜日から金曜日までの毎日、合計4-6時間行った。参加は、自発的なものであり、強制ではなかった。
  8. 主なアウトカム評価項目  
作業能力、コミュニケーション能力、自閉症状の改善、行動の改善、基本的な生活能力の主要5項目で、3段階の尺度（いつも良い、ときどき良い、だいたい悪い）とした。
  9. 主な結果  
作業能力、コミュニケーション能力、自閉症状の改善、行動の改善、基本的な生活能力は、3年間の介入後、有意に改善した。
  10. 結論  
森林での活動やレクリエーションは、自閉症者に対する重要な治療方法になる潜在性を有する。

表7 エビデンス・テーブル（4）：引用文献17

1. 研究題目  
活動前の疲労度別に見たスポーツ活動の効果について
2. 代表著者名（出版年）  
服部伸一（2001）
3. 目的  
活動前の疲労スコアの訴えレベルにも焦点をあて、対象者の疲労度別にスポーツ活動による効果をみる。
4. 研究デザイン  
対照群のない介入試験（ただし、疲労感の有訴の程度に基づき3群に層化して分析）
5. セッティング  
記載なし。
6. 対象者  
女子大学生48名（平均年齢18.1歳）で、のべ279名分のデータを採用した。
7. 介入内容  
6種目（バドミントン、プリズンボール、バレーボール、ソフトバレーボール、インディアカ、バドテニス）を1人1日1種目ずつ、日替わりで20分間実施した。
8. 主なアウトカム評価項目  
「疲労自覚症状調査」：30項目、それぞれの項目において、症状がない0点、少しある1点、かなりある2点、顕著にみられる3点、で評価した。  
「レクスコア」：10項目、それぞれの項目において、大変悪い-2点、やや悪い-1点、変化なし0点、少し良い1点、大変良い2点、で評価した。
9. 主な結果  
活動前の疲労スコア別の3群において、活動後のレクスコアが1よりも大きくなった項目数は、「疲労スコア0」群では7項目、「疲労スコア1-10」群では6項目、「疲労スコア11以上」群では4項目であった。
10. 結論 [注]  
対象者が活動前に疲労感が皆無という状態でスポーツ活動を行うと、疲労度は低く維持されたままでレクスコアは高くなり、より快適な状態になる。

[注] 論文中に単純で明快な結論が記載されていないため、筆者が要約・記載した。

表 8 エビデンス・テーブル (横断研究) : 引用文献 18

1. 研究題目	大学生のレジャーにおける退屈感
2. 代表著者名 (出版年)	田中節芳 (1999)
3. 目的	大学生のレジャーにおける心的状況を否定的感情、とくに退屈感に着目し、その基礎的な資料を得る。
4. 研究デザイン	横断研究
5. セッティング	1997年6月下旬に、無記名による質問紙調査を実施した。基礎教育科目として実施されている「スポーツ実習」および「体育実技」の時間中に行った。
6. 対象者	広島県内の2大学の学生650名を対象とし、有効回答数は585であった。男子67.5%、女子32.5%、1年生39.7%、2年生45.8%、3年生12.3%、4年生2.2%であった。
7. 主なアウトカム評価項目	7段階尺度からなる「レジャー能力」、16の設問で各1-5点で評価する「Leisure Boredom Scale (LBS)得点」、「1週間あたりのレジャー時間」、5段階尺度からなる「レジャーの適量感」、5段階尺度からなる「レジャーの満足感」
8. 主な結果	調査対象である学生の56.8%が、レジャーにおいて何らかの退屈感を感じ、そのうちの2.7%は頻繁に退屈間を感じていた。レジャー能力の自己評価が低い学生は、レジャー能力の自己評価が高い学生より有意にLBS得点が高かった。レジャー時間の適量感を感じる学生は、過多感を覚える学生より有意にLBS得点が高かった。単独でレジャーを過ごす学生は、集団で過ごす学生より、有意にLBS得点が高かった。レジャーにおける行動決定を他人に依存する学生は自分で決定する学生より有意にLBS得点が高かった。孤独感を強く感じる学生は、孤独感をあまり感じない学生より有意にLBS得点が高かった。
9. 結論	レジャー能力の自己評価、レジャーにおける集団の状況、性格の自己評価、レジャーにおける孤独感の有無、生活における時間意識の5つの要因とレジャーにおける退屈感との関連性が示唆された。

表9は、TREND 声明チェックリストに基づく、質評価の結果である。4編中1編はnRCTのため、このチェックリストに合致するが、他の3編は対照群が設定されていない。しかし、主に割付に関連する以外の部分では、適応できる項目がほとんどであるため、同リストを用いた。

方法においては、参加者の募集の詳細な説明、アウトカムではその精度(系統誤差を回避)を高めるための努力や工夫の記述がされていなかった。症例数(サンプルサイズ)の根拠や介入実施者、アウトカム評価者についての盲検化(ブラインド化)の情報が記載されている論文もなかった。欠損値データの処理や統計ソフトウェアの記載もなかった。

結果においては、フローチャートでの参加者の流れ、登録・完遂・脱落の情報が記載されていない論文がほとんどであった。またベースラインでのデータ数と、介入が終了するときのデータ数が違う場合の解析方法(ITT分析とOn treatment分析の両者)を行っている研究はなかった。結果の解釈において、交絡因子や測定誤差、潜在的なバイアスを踏まえて慎重な考察や一般化可能性(外的妥当性)を述べている論文もなかった。

表10は、Crで行われた唯一の論文のSTROBE 声明チェックリストに基づく質評価である。目的では、多重仮説を検証するという記載であり、特定の仮説を含んだ明確なものではなかった。方法では、参加者の適格基準、バイアスへの対応、研

究サイズの根拠、統計・分析の詳細な説明の不足、結果では、フローチャートや非回答者の理由が記載されていない、また交絡因子の調整がなされていない。結果の総合的な解釈や一般可能性についての記載は、多重仮説の設定があったため十分ではなかった。

表 11 は、表 9-10 のチェックリストに含まれていない基本事項で、実施・記載がないと論文の質に大きなダメージを与える諸点である。多重仮説

の論文が 40% (2/5)、レジャー・レクリエーションの明確な定義づけのない論文が 100% (5/5)、インフォームド・コンセントについての記載のない論文が 100% (5/5)、倫理審査委員会の承認の有無の記載がない論文が 100% (4/4)、有害事象の記載がない論文が 100% (4/4) であった。参考文献 (引用文献) において、査読を要しない学会発表抄録や図書などが、引用の半分以上だった論文が 60% (3/5) であった。

表 9 TREND 声明に基づく nRCT の質評価\* (引用文献 7, 8 より一部改変作表)

[分子：記載・実施ありの論文数 分母：対象となる論文数 -：該当論文なし] で表記

	no	内容	記載状況
タイトル・抄録	1	介入の単位の割振りをどのように行ったかの方法の情報。	0/1(0%)
	2	構造化抄録を推奨。	0/4(0%)
	3	対象となる母集団や研究標本についての情報。	3/4(75%)
はじめに 背景	4	科学的背景と合理的根拠の説明。	3/4(75%)
	5	行動的介入を計画する際に用いられた理論。	2/4(50%)
方法			
参加者	6	参加者の適格条件。 さまざまなレベルにおける募集と抽出計画を含む。	0/4(0%)
	7	募集方法 (紹介、自己選択など)、抽出が体系的に行われた場合には抽出方法も含める。	1/4(25%)
	8	募集のセッティング。	0/4(0%)
介入	9	データが収集されたセッティングと場所。 介入についての詳細、それが実際にいつどのように行われたか。	3/4(75%)
	10	内容：何を与えたか。	4/4(100%)
	11	介入方法：どのように与えたか。	4/4(100%)
	12	介入単位：参加者はどのようにグループ分けされたか。	1/1(100%)
	13	介入者：誰が介入したか。	3/4(75%)
	14	セッティング：介入はどこで行われたか。	3/4(75%)
	15	介入の量。	4/4(100%)
	16	期間：各個人にどれくらいの期間行う予定だったのか。	4/4(100%)
	17	コンプライアンスや研究参加継続のための措置(報酬等)。	0/4(0%)
目的	18	特定の目的と仮説。	4/4(100%)
アウトカム	19	明確に定義された主要・二次的アウトカム評価項目。	4/4(100%)
	20	データ収集に用いられた方法および測定を向上させるために用いたあらゆる方法。	0/4(0%)
	21	心理測定学および生物測定学的特性などの検証済みの手法についての情報。	3/4(75%)
症例数	22	どのように目標症例数が決定されたか。あてはまる場合は、中間解析と中止基準の説明。	0/4(0%)
割付方法	23	割付の単位。	1/1(100%)

表 9 : 続き

	24	研究状態への割付に用いた方法、制限の詳細を含む。	1/1(100%)
	25	非ランダム化による潜在的なバイアスを最小限にするために行った措置。	0/1(0%)
ブラインド化/ マスキング	26	参加者、介入実施者、アウトカム評価者に対し研究状態の割付が盲検化されたか。どのように盲検化を達成し、評価したかについての説明。	0/4(0%)
解析の単位	27	介入の有効性を評価するために解析された単位についての記述（個人、集団、地域など）。	2/4(50%)
	28	解析の単位が割付の単位と異なる場合、これを説明するために用いた解析方法から推定された標準誤差推定値の調整や多重解析など。	—
統計的手法	29	主要アウトカムの群間比較に用いられた統計的手法。	1/1(100%)
	30	サブグループ解析や調整解析などの付加的な解析手法。	—
	31	欠損データへの対処法（用いられた場合）。	0/4(0%)
	32	使用された統計ソフトウェアやプログラム。	0/4(0%)
<b>結 果</b>			
参加者の流れ	33	各段階における参加者の流れ：登録、割付、割付と介入曝露、追跡、解析（フローチャートを強く推奨）。	0/4(0%)
	34	登録：適格基準でスクリーニングされた人数、適格または不適格・登録を拒否した人数、登録人数。	0/4(0%)
	35	割付：割付られた人数。	1/1(100%)
	36	割付と介入曝露：各研究条件に割付られた人数および介入を受けた人数。	1/1(100%)
	37	追跡：各研究状態ごとの追跡完了者数、不完了者数。	0/4(0%)
	38	解析：各研究状態ごとの主要解析の対象者数、除外者数。	4/4(100%)
	39	研究プロトコルからの逸脱とその理由も理由。	0/4(0%)
募 集	40	募集期間と追跡期間を決特定する日付	
ベースライン データ	41	各研究状態のベースラインにおける人口統計的、臨床的な特徴。	1/4(25%)
	42	各研究状態のベースラインにおける特徴のうち、特定の疾患予防研究に関連するもの。	1/1(100%)
	43	脱落例と追跡例のベースラインでの比較を患者、全体と各研究状態の患者ごとに示す。	0/4(0%)
	44	ベースライン時の試験対象者と目標母集団との比較。	0/4(0%)
ベースラインに おける同等性	45	ベースラインにおける群間の同等性を示すデータ、ベースラインにおける差異を調整するために用いた統計的手法。	1/1(100%)
解析された数	46	各解析における各研究状態の人数（分母）を示し、可能であれば、結果の記述には絶対数を用いる。	3/4(75%)
	47	Intention-to-treat (ITT) 解析を行ったかどうかを示し、もし ITT 解析でない場合は、逸脱者の解析をどのように行った記述。	0/4(0%)
アウトカムと 推 定	48	主要・副次アウトカムのそれぞれについて、各研究状態の結果を要約し、推定されたエフェクトサイズと精度の信頼区間を示す。	1/1(100%)
	49	無効および否定的な知見を記載する。	2/4(50%)
	50	介入がたどるべき経過として、事前に特定された因果経路について検証された場合には、その結果を記載する。	1/4(25%)
補助的分析	51	サブグループ解析や制限解析など、実施されたその他の解析についての概略。その分析が予定されたものか探索的なものかを示す。	1/1(100%)
有害事象	52	各研究状態におけるすべての重要な有害事象や予期せぬ効果の概略（要	0/4(0%)

表9：続き

考 察	約した指標、推定されたエフェクトサイズ、信頼区間)。	
解 釈	53	結果の解釈は、研究仮説、潜在的なバイアスの要因、測定誤差、多重性 解析、研究における限界や欠点を考慮して記載。 0/4(0%)
	54	結果の考察は、介入の作用として意図されていたメカニズム（因果経路） またはそれに代わるメカニズムや説明を考慮して行う。 1/4(25%)
	55	介入の成果および介入実施の障壁についての考察、介入がどれだけ忠実 に実施されたか。 0/4(0%)
	56	研究面、プログラム面、政策面に対する影響。 1/4(25%)
一般化可能性	57	研究集団、介入の特性、追跡期間、インセンティブ、コンプライアンス、 研究に関わる特性の場所やセッティング、状況的な問題を考慮した試験 結果の一般化可能性（外的妥当性）。 0/4(0%)
全体としての エビデンス	58	現時点におけるエビデンスや理論を考慮とした包括的解釈。 3/4(75%)

[注] 多くの行動的介入においては、参加者または介入実施者のマスキング（盲検化）が必要でない場合や不可能な場合がある。介入を計画する際に用いた理論（項目5）については、「介入」項目では10-17の一部として報告することも可能である。試験対象者と目標母集団のベースラインにおける比較（項目41-44）は、項目57の一部として報告することも可能である。

\* 分母の対象論文数が1となる部分は、対照群を設定してあるのが1編だけであるためである。

表10 STROBE 声明に基づく横断研究の質評価（引用文献9,10 から一部改変作表）

「記載あり：1 記載なし・実施していない：0 該当しない項目：-」で表記

	no	内 容	記載状況 N=1
タイトル・抄録	1	タイトルまたは抄録のなかで試験デザインを一般に用いられる用語 で明示する。	0
	2	抄録では、研究で行われたことと明らかにされたことについて、十 分な情報を含み、かつバランスのよい要約を記載する。	1
はじめに			
科学的背景/理論的根拠	3	研究の科学的な背景と論拠根拠を説明する。	1
目 的	4	特定の仮説を含む目的を明記する。	0
方 法			
研究デザイン	5	研究デザインの重要な要素を論文のはじめで示す。	0
セッティング	6	セッティング、実施場所のほか、基準となる日付については、登録、 曝露、追跡、データ収集期間を含めて明記する。	1
参加者	7	適格基準、参加者の母集団、選択方法を示す。	0
変 数	8	すべてのアウトカム、曝露、予測因子、潜在的交絡因子、潜在的な 効果修飾因子を明確に定義する。該当する場合は、診断方法を示す。	0
データ源/測定方法	9	関連する各因子に対して、データ源、測定・評価方法の詳細を示す。 2つ以上の群がある場合は、測定方法の比較可能性を明記する。	1
バイアス	10	潜在的なバイアス源に対応するためにとられた措置があれば潜在的な バイアス源に対応するためにとられた措置があればすべて示す。	0
研究サイズ	11	研究サイズがどのように算出されたかを説明する。	0
量的変数	12	量的変数の分析方法を説明する。該当する場合は、どのグルーピン	1

表 10 : 続き

		グがなぜ選ばれたかを記載する。	
統計・分析方法	13	交絡因子の調整に用いた方法を含め、すべての統計的方法を示す。	0
	14	サブグループと相互作用の検証に用いたすべての方法を示す。	1
	15	欠損データをどのように扱ったかを説明する。	0
	16	該当する場合は、サンプリング戦略を考慮した分析法について記述する。	—
	17	あらゆる感度分析の方法を示す。	—
<b>結 果</b>			
参加者	18	研究の各段階における人数を示す。	1
	19	各段階での非参加者の理由を示す。	0
	20	フローチャートによる記載を考慮する。	0
記述的データ	21	参加者の特徴（例：人口統計学的、臨床的、社会的特徴）と曝露や潜在的交絡因子の情報を示す。	0
	22	それぞれの変数について、データが欠損した参加者数を記載する。	1
アウトカムデータ	23	アウトカム事象の発生数または集約尺度を示す。	1
主要な結果	24	調整前の推定値と該当する場合は交絡因子での調整後の推定値、そしてそれらの精度（例：95%信頼区間）を記述する。どの交絡因子がなぜ調整されたかを明確にする。	0
	25	連続変数がカテゴリー化されているときは、カテゴリー境界を報告する。	1
	26	意味がある場合は、相対リスクを意味をもつ期間の絶対リスクに換算することを考慮する。	—
他の解析	27	その他に行われたすべての分析（例：サブグループと相互作用の解析や感度分析）の結果を報告する。	1
<b>考 察</b>			
鍵となる結果	28	研究目的に関しての鍵となる結果を要約する。	1
限 界	29	潜在的なバイアスや精度の問題を考慮して、研究の限界を議論する。潜在的バイアスの方向性と大きさを議論する。	1
解 釈	30	目的、限界、解析の多重性、同様の研究で得られた結果やその他の関連するエビデンスを考慮し、慎重で総合的な結果の解釈を記載する	0
一般化可能性	31	研究結果の一般化可能性(外的妥当性)を議論する。	0
<b>その他の情報</b>			
研究の財源	32	研究の資金、本研究における資金提供者の役割を示す。該当する場合には、現在の研究の元となる研究についても同様に示す。	—

[注] 横断研究において、該当する場合には曝露群と非曝露群に分けて情報を記述する。

表 11 研究の質にダメージを与える事項

項 目	該当論文数(%)	理 由
1. 研究の定義づけ		
(1) 多重仮説の設定あり	2/5(40%)	・論文が冗長になる。(一仮説・一研究の原則) ・交絡因子を調整した分析でないと、不適切な結果を導き出す。
(2) レジャー・レクリエーションの明確な定義づけなし	5/5(100%)	・スポーツ、身体活動などとは異なることを示さないと、学会外の読者はレジャーやレクリエーションの効果とは判断

表 11: 続き

		しない。
2. 倫理・安全面への配慮*		
(1) インフォームド・コンセントについての記載なし	5/5(100%)	・ キャンプ等に乗るようなコンビニエンス・データでも、当該目的以外の研究のために使用する場合には、インフォームド・コンセントは必須である。
(2) 倫理審査委員会の承認の記載なし	4/4(100%)	・ 介入研究、とくに行動を限定させたり、侵襲性を伴う研究では必須である。 ・ 前述の研究の場合、倫理審査委員会の承認を得ていない論文は、学術雑誌は受理してはならない。 ・ 介入研究で、侵襲性が高いものは、国の臨床試験登録 (UMIN-CTR) ** に事前に登録してから実施しなければならない。
(3) 有害事象の記載なし	4/4(100%)	・ とくに介入研究において生じた健康被害や事故などの有無を必ず記載しなければならない。 ・ 「その介入による利益は害より大きいか?」 <sup>19) 20)</sup> の重要な判断 (Judgment) 材料となる。
3. 参考文献		
出典の半数以上が査読のない文献	3/5(60%)	・ 査読のない文献 (学会抄録、図書など) の引用が多いことは、先行研究の事前調査や考察・解釈の信頼性に影響を及ぼす可能性がある。

[注]\* 国 (文部科学省・厚生労働省) の「疫学研究に関する倫理指針」(平成 19 年 8 月 16 日全部改正) を理解してから研究を実施すべきである。

\*\* 正式名称：大学病院医療情報ネットワーク研究センター (University Hospital Medical Information Network: UMIN)、臨床試験登録：Clinical Trial Registry: CTR) である。

#### 4. 考察

過去 15 年間のレジャー・レク誌において、5 編 (11.9%) の研究が疫学的な研究手法を用いていた。その中で、介入研究である 4 編<sup>14) -17)</sup> は、独自の介入方法を用いており興味深い内容であった。観察研究の中の横断研究 1 編<sup>18)</sup> は、あえてネガティブなイメージを題目とした新規的な課題であった。

しかしながら、自然科学、とくに人を対象とした教育や治療効果などを実証するには方法論的な問題点も多いことが、チェックリストとの照合から明らかになった。以下にそのウィーク・ポイントを示しつつ、論文の質向上のための対策を議論する。

##### (1) タイトル・抄録

構造化抄録を用いている介入研究は 0 であり、研究デザインを明記しているものもなかった。構造化抄録を採用する学術雑誌が急増している背景とともに、研究者が短時間で的確に研究内容を理解できるため、疫学を扱うレジャー・レクリエー

ション研究でも構造化抄録の導入が推奨される。

##### (2) 参加者 (対象者)

参加者の募集方法や適格基準 (除外基準も含む) の説明や対象母集団からの抽出の記載がないため、選択バイアスの可能性やその影響の大きさが強く疑われることになる。こうした問題を生じさせないためにも、事前にチェックリストの項目に沿っての記載が有効である。合わせて、それらをフローチャートで書き示すことが常套手段となっている。

また、ドロップアウトの人数や拒否者数、その理由などもエビデンスの構築には不可欠であるが、詳細な記載がほとんどなされていなかった。大多数には受け入れられる教育・介入や調査方法であっても、少数には拒否者や、ついて行けない者が存在する事実も含めて、その介入効果を判断する必要があるからである。さらには、継続した者 (群) の前後比較だけでなく、ドロップアウトした者も含めての解析 (ITT 解析) も正しい評価として必須である。

### (3) アウトカム項目

回答・測定結果の質を高めるための工夫や努力が記載されていなかった。この具体例として、「pre/post のアンケートを行うのに、同じ時間帯に、外乱（周囲の話し声など）が及ばない静かな部屋で行った。」とか「ウエスト囲の測定は、研究のために依頼した第三機関の熟練した検者が、盲検化され、介入群か対照群かわからないようにして実施した。」などの事項が該当する。これにより情報バイアスを減少させることができるからである。

### (4) 解析

欠損データの取り扱いや統計ソフトウェアの記載がなかった。介入研究においては、脱落者の評価も含めて行うことが必須（ITT 解析）であり、最後まで残った者だけの前後比較（Per-Protocol 解析）だけでは、介入プログラムの遵守の程度を無視することになる。つまり、「都合の良いところのデータ取り」という問題が生じる。

### (5) 解釈

交絡因子や潜在的なバイアス、測定誤差などの研究の限界と先行研究の結果、さらには因果経路（メカニズム）を踏まえての慎重な記載が必要であるが、いずれの研究も十分ではなかった。これは、全体としてのエビデンスにも共通する。チェックリストにおける項目にしたがって論述することは、直接的に客観的な自己の論文批評の助けになる。

### (6) 一般化可能性（外的妥当性）

一般化可能性は、チェックリスト内の前述項目、つまり内的妥当性の精度が高くなければ意味を持たないが、対象としている疫学的論文においては論述がなかった。質を高めるためには、まず内的妥当性が重要であるものの、人を対象としている研究であるならば、研究結果がどのように活用されうるかの方向性を示す必要がある。

### (7) 研究の質にダメージを与える他の事項

冗長で混沌とした論文記載を防ぐためにも、多重仮説は回避すべきである。1 仮説・1 研究を原則とし、説明変数（独立変数）として、いくつかの項目がある場合には、解析において調整解析を行うことが必要であり、それが不可能であるならば、当初から 2 研究として別個に行うべきである。

倫理面への配慮として、基本的なインフォームド・コンセントについての記載もなかった。授業中にアンケートに回答してもらう程度のコンビニエンス・データであっても、研究目的で使用するならば必須である。また、人に対して一定の行動制約や行動変容、なんらかの曝露を強いる、いわゆる介入研究では、インフォームド・コンセントだけでは不十分であり、実施の可否は事前に第三機関の判断に委ね、承認を得てから行う必要がある。

さらに有害事象については、事故がなかったことも含めて明記すべきである。介入の総合的な効果を判断するには、「利益は害より大きいこと」を示せることが大前提であるからである。具体例を挙げれば、99 名には大成功のスノーボードを中心とした冬山キャンプでも、1 名の死亡者を出してしまうようなプログラムでは、推奨できないということである。

参考文献では、総じて図書や学会抄録集、学位論文（修士）などからの引用が多く、国内外の学会誌からの引用が少ないことも特徴的であった。

### (8) 論文の質を高めるための必要条件

1993-2007 年までのレジャー・レク誌に掲載された人を対象とした疫学的論文のレビューから、研究の質を高めるための 4 つの必要条件を図 1 に示した。1 点目として、データベースやプラットフォームを活用した先行研究、とくに質の高い論文の入念な調査が必要である。2 点目として、核心部分であるレジャー活動やレクリエーションの用語の定義が不可欠である。3 点目は、人を対象とした疫学的研究の倫理面に対する正しい理解が必要である。4 点目は、著名な各種の声明チェックリストに応じた、あるいは参考にした研究計画・実施・論文記載が必要である。チェックリストに即応した記載は、結果として、論文の質向上につながる著効性があるため強く推奨したい。

### (9) 本研究の限界

本研究には、いくつかの限界が含まれている。1 点目は、系統的な文献収集を行っているが、レジャー・レク誌に限定しているため、国内外の他の雑誌に nRCT や Os が発表されている可能性が高いという、選択バイアスの問題を含んでいる。今後、それらを包括したシステムティック・レビ

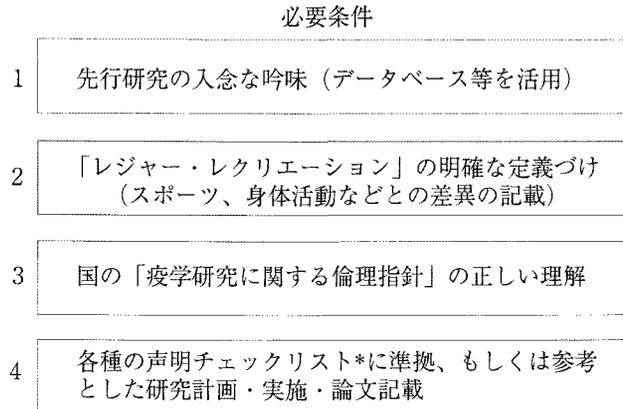


図1 人を対象としたレジャー・レクリエーション研究の質を高めるための4つの必要条件

ューが望まれる。

2点目は、論文には記載されていないが、実際には実施している項目もあるかもしれない、著者への問い合わせを行っていないため確かな事実はいえないということである。システムティック・レビューのチェックリストにおいて、RCTでは「QUOROM 声明チェックリスト」<sup>21)</sup>、Osでは「MOOSE チェックリスト」<sup>22)</sup>があるが、ともにチェック項目の中に記載のない項目がある場合には、著者に直接問い合わせるべきことが示されている。また、Soaresら<sup>23)</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RCT58編を分析した結果、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、必ずしも実施したことすべてを掲載しているとは言えず、臨床研究登録（Clinical Trial Registry: CTR）の研究計画書を調べたり、研究者に直接問い合わせるとともに、予防のためにもチェックリストに基づいて最初から論文を記載すべきことを述べている。

3点目としては、レジャー・レクリエーション研究において、特徴的あるいは独自の意義ある項目があるかもしれないものの、本研究では明らかにできていないことが挙げられる。探索的な研究デザインとして、それに焦点を当てた研究が期待される。

1993-2007年までのレジャー・レク誌に掲載された人を対象とした疫学的論文のレビューから、論文の質を高めるために、次の4点の必要条件を整理できた。(1)各種のデータベースなどを活用した国内外の先行研究、とくに質の高い論文の入念な調査、(2)核心部分であるレジャー活動やレクリエーションの用語の定義、(3)倫理面への配慮と適した事前措置、(4)国際的にコンセンサスの得られている声明チェックリストに応じた、あるいは参考にした研究計画・実施・論文記載、である。

## 附記

文献の網羅的な収集と内容の吟味に関して、関東学院大学の鈴木秀雄教授の多大なご支援を賜りました。この場をお借りして深謝いたします。

なお、本研究は、平成19年度厚生労働省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（H18-循環器等（生習）一般-036）」の一部として実施した。

## 註

本論では、「evidence（エビデンス）」の定義を「科学的根拠」とし、レジャー活動とレクリエーションにおける「効果の科学的根拠」を考究する。

また、「レジャー活動」と「レクリエーション」の

## 5. 結論

研究分野の範疇としては、効果を定量化することが可能な研究すべてとなるが、レジャー・レクリエーション史や哲学、人類学などの人文分野やレクリエーションの具体的な指導・方法論を考究する分野は含めない。

## 引用文献

- 1) キャンベル共同計画：アクセス、  
<http://www.campbell.gse.upenn.edu/>
- 2) 財団法人厚生統計協会、国民衛生の動向、12、2004
- 3) 上岡洋晴・津谷喜一郎・本多卓也ら、「レジャー活動」と「レクリエーション」に関するランダム化比較試験のシステムティック・レビュー、レジャー・レクリエーション研究 60：29-37、2008
- 4) David Moher, Kenneth F Schulz, Douglas G Altman, et al., The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials, JAMA 285：1987-1991, 2001
- 5) 津谷喜一郎・小島千枝・中山健夫訳・解説、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、2002
- 6) 黒田信寛、就任時の学会を振り返って、レジャー・レクリエーション学会の歩み 1964-1995 32：6、1995
- 7) Don C Jarlais, Cynthia Lyles, Nicole Crepez, et al., Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND Statement, Am J Public Health, 94:361-366, 2004
- 8) 中山健夫(訳)：行動的介入および公衆衛生的介入を評価した非ランダム化研究報告の質の改善：TREND 声明。(中山健夫,津谷喜一郎編「臨床研究と疫学研究のための国際的ルール集」ライフサイエンス出版、東京)、194-201、2008
- 9) Eric Elm, Douglas Altman, Matthias Egger, et al., The strengthening the reporting of observational studies in epidemiology (STROBE) Statement: guidelines for reporting observational studies, Ann Inter Med, 147:573-577, 2007
- 10) 上岡洋晴・津谷喜一郎(訳)：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(中山健夫, 津谷喜一郎編、臨床研究と疫学研究のための国際的ルール集、ライフサイエンス出版、東京)、202-209、2008
- 11) 高橋美絵・上岡洋晴・津谷喜一郎ら、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ムティック・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、2007
- 12) Isabelle Boutron, David Moher, Peter Tugwell, et al., A checklist to evaluate a report of a nonpharmacological trial (CLEAR NPT) was developed using consensus, J Clin Epidemiol 58:1233-1240, 2005
- 13) 上岡洋晴・津谷喜一郎・川野因ら、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、農学集報 53：81-89、2008
- 14) 岡村泰斗・飯田稔・星野敏男ら、環境教育プログラムを導入したキャンプの効果、レクリエーション研究 33：1-6、1996
- 15) 関智子・飯田稔・橋直隆ら、キャンプカウンセラーの性役割がキャンパーの性役割意識に及ぼす影響、レクリエーション研究 33：17-23、1996
- 16) Iwao Uehara and Kenji Sasaki, The possibility of forest activities in the autistic disabilities treatment by utilizing the rural forest, J Leisure Recreation Studies 40:59-67, 1999
- 17) 服部伸一・前橋明、活動前の疲労度別にみたスポーツ活動の効果について、レクリエーション研究 44：11-18、2001
- 18) 田口節芳・富永徳幸・折本浩一ら、大学生のレジャーにおける退屈感、レクリエーション研究 40：11-23、1999
- 19) David Atkins, Dana Best, Peter A Briss, et al., Grading quality of evidence and strength of recommendations, BMJ328:1490-1497,2004

- 20) 津谷喜一郎・中山健夫・島村治子(訳)、エビデンスの質とお勧め度のグレーディング、薬理と臨床 32 : 1241-1254、2005
- 21) David Moher, Deborah J Cook, Susan Eastwood, et al., Improving the quality of reports of meta-analyses of randomized controlled trials: the QUOROM statement, *Lancet* 354:1896-1900, 1999
- 22) Donna F Stroup, Jesse A Berlin, Sally C Morton, et al., Meta-analysis of observational studies in epidemiology, *JAMA*, 283:2008-2012, 2000
- 23) Heloisa P Soares, Stephanie Daniels, Ambuj Kumar, et al., 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, 2004

( 受付：2008年5月26日 )  
( 受理：2009年1月29日 )