

Summary

The purpose of this study was to clarify the relationship between frequency of bathing in a spa and at home and the characteristics of the body, blood, and lifestyle of white-collar and blue-collar male employees.

To recruit subjects, two baseline data for randomized controlled trials in two places in Shimane Prefecture and in Nagano Prefecture were used. In the two-month period between August and September 2006, 43 of 311 white-collar male employees aged between 30 and 57 years in the Unnan municipal office volunteered to participate in this study. Similarly, 44 blue-collar workers in Nagano Prefecture volunteered to be involved in this study as a result of an appeal for volunteers in local newspapers and public information journals published by large local enterprises from September through November 2006. A total of 87 men were subjects of this study. Items evaluated in this study with respect to physique included height, weight, body mass index, waist circumference, hip circumference, and percent of body fat. With respect to strength, they were grip strength, abdominal strength, back strength, and anteflexion. With respect to characteristics of the blood, they were serum glucose, hemoglobin A1c, fructosamine, lactic acid, total cholesterol, triglyceride, HDL cholesterol, LDL cholesterol, free fatty acid, GOT, GPT, γ -GTP, cholinesterase, natural killer cell activity, T cell, B cell, CD4+, CD8+, CD4/8 and uric acid. With respect to mood, Profile of Mood States was evaluated. With respect to lifestyle, the frequency of monthly bathing in a spa and at home was evaluated.

No significant correlation was revealed between frequency of monthly spa bathing and the health indices in the white and blue-collar male workers. The items that were significantly correlated ($p < 0.05$) with frequency of monthly bathing at home were GOT (positive), CD8+ (positive), and mental stress (negative).

In this study, frequency of monthly spa bathing was found to have no significant correlation with the health indices.

総説

温泉に関する研究の質を高めるための
チェックリストや声明の活用の意義：
疫学・臨床研究のエビデンス・グレーディングと
研究デザイン

上岡洋晴¹⁾、津谷喜一郎²⁾、高橋美絵³⁾、岡田真平³⁾、塩澤信良⁴⁾

1) 東京農業大学地域環境科学部

2) 東京大学大学院薬学系研究科医薬政策学講座

3) 身体教育医学研究所

4) 岡山大学大学院医歯薬学総合研究科老年医学分野

**Significance of Utilizing Checklists and Statements
for Improvement of the Quality of Studies
Concerning Hot Springs:
evidence grading and various study designs of
epidemiological and clinical studies**

Hiroharu KAMIOKA¹⁾, Kiichiro TSUTANI²⁾, Mie TAKAHASHI³⁾, Shinpei OKADA³⁾, Nobuyoshi SHIOZAWA⁴⁾

1) Faculty of Regional Environment Science, Tokyo University of Agriculture

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

3) Physical Education and Medicine Research Center, Mimaki Social Welfare Corporation

4) Department of Longevity and Social Medicine, Okayama University
Graduate School of Medicine, Dentistry and Pharmaceutical Sciences

Summary

The purpose of this study was to review several kinds of checklists and statements of studies in human subjects, and to examine reporting methods in order to improve the quality of studies concerning hot springs.

For randomized controlled trials (RCTs), "the CONSORT Statement" (Moher et al. *JAMA* 2001) consisting of 22 terms is the most well known checklist. In addition, different versions and new checklists have been developed depending upon the content of the studies (such as intervention methods) or field of the studies. There is also "the QUOROM Statement" (Moher et al. *Lancet* 1999) for a systematic review (SR) of RCTs, "the TREND Statement" (Jarlais et al. *Am J Public Health* 2004) for nonrandomized controlled trials, "the STROBE Statement" (Elm et al. *Ann Intern Med* 2007) for observational studies, and "the MOOSE Checklist" (Stroup et al. *JAMA* 2000) as a SR of observational studies. With regard to studies on hot springs, however, terms in the checklists and the state-

ments described above are insufficient or inappropriate because of difficulties in performing such studies in a blinded manner, and the diversity of intervention methods, such as the quellen charakter.

Improvement of the quality of reports is important for validation of evidence. In order to improve the quality of assessments and reports of studies on hot springs, it may be necessary to develop a unique checklist based on the above-described statements and checklists.

Key words : Balneotherapy, Clinical Trials, Grading of Evidence, Recommendation, Checklists

1 緒言

1990年代後半以降、日本においても「エビデンスに基づいた医療」(evidence based medicine: EBM)、さらに「エビデンスに基づいた健康政策」(evidence based health policy: EBHP)という言葉が用いられるようになった。

臨床領域で目の前の患者についての意志決定に関するものがEBMで、社会領域でマスを対象にしたそれがEBHPであり、後者はエビデンスに基づくヘルスケア (evidence based health care: EBHC) と称されることもある。

EBMにおいてもEBHPにおいても、そこではグレーディングの概念が用いられる。厚生労働省も支持¹⁾しているこうしたEBMやEBHPの核心的な考え方の正しい理解が必要である (Table 1)。

温泉に関する研究においても同様である。温泉は「介入」のひとつとみなされ、グレーディングの概念が導入される。この質のグレーディング・スケールは、エビデンスの強さ (strength of evidence) について、「お勧め度」 (strength of recommendation) についての2種類があり、それぞれ複数存在するが、前者は基本的には研究デザインによる。そして、いくつかの研究デザインごとに、論文中に記載すべき、あるいは研究で実施すべき事項を定めたチェックリストや声明が、著名な雑誌などで紹介されている。これらは、学術雑誌に投稿される論文の質を高め、エビデンスを正しく「つくる」「つたえる」ためである。

ランダム化比較試験 (randomized controlled

trial: RCT)については、「改訂版CONSORT声明^{2),3)}が良く知られている。世界的な*Ann Inter Med*, *BMJ*, *JAMA*, *New Engl J Med*, *Lancet*などの11の雑誌からなる「国際医学雑誌編集者委員会」(International Committee of Medical Journal Editors: ICMJE) の「生物医学雑誌への統一投稿規定」においても、RCT論文を投稿する際には、改訂版CONSORT声明に基づいているかのチェックリストのシートがあり、それを完成させ、付しておかないと受け付けられないシステムが構築されている。こうしたチェックシステムを採用する学術雑誌は増えている。論文の質を保証するために、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、さらには研究の質そのものを確保するための予防措置である。

Moherら⁴⁾は、1996年の最初のCONSORT声明⁵⁾が出される前後において、掲載された論文の質を比較したところ、声明後の論文の方が質が高まっていることを報告している。

津谷⁶⁾は、Fig.1に示すようにエビデンスには3段階があり、「つくる」部分が臨床研究、「つたえる」部分がシステムティック・レビュー (メタ・アナリシス)、「つかう」部分が臨床医やガイドライン作成者などの現場に従事する者たちであることを示している。

世界で報告されるRCTの質が高まることは、「エビデンスをつたえる部分のシステムティック・レビュー」が、より正しく真実を述べることを意味している。一方で、温泉研究においては、RCTが難しい場合が多く、非ランダム化比

Table 1 エビデンス・グレーディング
(治療の有効性に関して)¹⁾

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

較試験 (nRCT) や観察研究が主流を占めてきた。こうした研究デザインにおいても、正しい論文の記載や分析方法によって質を確保でき、エビデンスの構築に寄与するものと考えられる。

そこで本研究は、人を対象とした研究デザインについての各種のチェックリストや声明を整理し、温泉に関する研究の質を高めるための報告方法を考察することを目的とした。

II 方法

1990年代後半以降における疫学研究や臨床研究の評価や質を高めるためのチェックリストや声明の特徴を一覧表にした。研究の質に関して議論となっている諸点を記述した。さらに各種の声明やチェックリストを例示しながら、温泉研究として実施すべき研究の方向性を考察した。

III チェックリストと声明

温泉研究の評価や質向上において、参考となるチェックリストと声明をTable 2-3に示した。Table 2は、RCTとそのシステマティック・レビュー (systematic review: SR) に関する一覧である。改訂版CONSORT声明がRCTのスタンダードとなるが、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ある。それぞれの領域・分野に応じて、改変していることが理解できる。

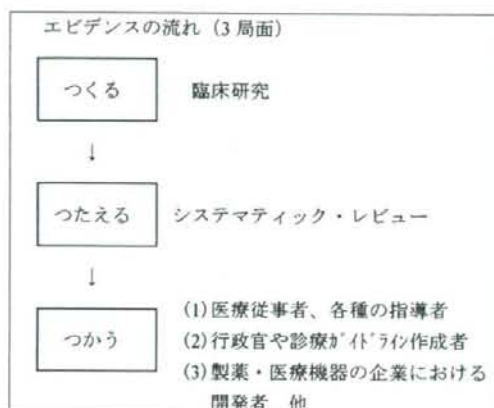


Fig.1 エビデンスの流れとシステマティック・レビュー
(津谷, 2003より作図)⁹⁾

Table 3は、nRCTと観察研究 (observational study: OS) およびSR (OS) に関する一覧である。過去の温泉研究の多くは、これらのデザインを用いた論文であり、今後も実施しやすいデザインであることから、温泉独自のチェックリストが望まれる。温泉研究では、RCTは多いとは言えず、現時点において、エビデンスを統合するためにメタ・アナリシスを含んだOSのSRが必要だと考えられる。

今回取り上げたチェックリストや声明は、デルフィ・コンセンサス法 (Delphi consensus method) が用いられることが多い。例えば、Boutronら¹²⁾ は、CLEAR NPTを開発するのにコクラン共同計画のメンバー41人、生物統計学者と疫学者55人、薬物以外の臨床研究実施者58人に研究協力の依頼をして、36%に相当する55人の専門家の参加の中で、チェックリストの絞り込み (当初38項目) を3ラウンド実施した。項目として必要な程度を満点10で各専門家に評価してもらい、その平均得点の高い順に残して策定している。

温泉研究においても、研究の質を高めるための独自のチェックリストを作成するのであれば、上記のように臨床医 (専門分野) や疫学・生物統計学者、その他学際領域の多くのエキスパートが参加して行うことが望ましいであろう。

Table 2 ランダム化比較試験(RCT)の質を高めるためのチェックリスト・声明と特徴及び温泉研究への応用性

研究デザイン	細目	名称または略称	特徴と温泉研究への応用性**
システマティックレビュー (SR:RCTによる)	RCTのSR	QUOROM声明 ⁷⁸⁾	RCTのシステマティックレビューを行う上での特有のチェック項目がある。例えば、出版バイアスについての項目やメタ・アナリシスでのデータ統合の方法や感度分析も含まれている。18項目で構成されている。
ランダム化比較試験 (RCT)	RCTを包括	改訂版CONSORT 声明 ²³⁾	RCTの基本となるチェックリストであり、広く医学分野においてコンセンサスが得られている。22項目で構成されている。温泉研究においても、基礎知識として重要である。
	クラスターRCT	Cluster RCTs ⁹⁾ *	クラスター(多施設間: multi-center)によるRCTのためのチェックリストで、改訂版CONSORT声明の拡張版(22項目)である。倫理問題も絡み、個人別の割付で実施が難しい状況において、コミュニティーや病院、施設等で割付をするので、温泉研究では、今後、数多くの研究が期待できる研究デザインである。
	非劣勢・同等性	Noninferiority and equivalence RCTs ¹⁰⁾ *	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証明するものであり、特殊性がある。こうした点を踏まえて、さらに改訂版CONSORT声明をアレンジした22項目である。温泉研究では、例えば、ある泉質で効果があると実証されているスタンダードな温泉と、新たに仮説として設定した別の泉質の温泉の効果に差がないことを明らかにするようなデザインが想定される。
	有害事象	Harms in RCTs ¹¹⁾ *	有害事象を明記することで、総合的にある介入の効果の評価することに役立て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。温泉研究では、湯あたりなどの入浴に伴う禁忌事項や入浴中の事故なども、このチェック項目とすれば含まれる。
	薬以外のRCT	CLEAR NPT ¹²⁾	薬以外の臨床試験ではブラインドが困難であったり、外科療法や理学療法などではプロバイダーのスキルの影響が大きいなどの特殊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。
	鍼治療	STRICTA ¹³⁾¹⁴⁾	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。鍼独自の項目が多く、温泉特有の介入手法を吟味する上で参考になる。
	理学療法	PE德罗スケール ¹⁵⁾¹⁶⁾	PE德罗の正式名称は、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り、このチェックリストは、同分野においては広く用いられている。内的妥当性10項目と、外的妥当性1項目から構成されている。
	温泉療法	PE德罗スケール改変版 ¹⁷⁾	原著の10項目に加えて、著者が3項目を追加した温泉研究のためのチェックリストである。その13項目で良いかについては、コンセンサスが得られていないという点で限界がある。
	健康づくり (生活・運動指導)	Cochrane レビュー(転倒予防 ¹⁸⁾ ・ 漸増的筋力増強運動 ¹⁹⁾ と PE德罗スケールの改変版 ²⁰⁾ *	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PE德罗スケールの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。コクラン・レビューとPE德罗スケールは重複している部分が多い。例えば、生活指導を温泉療法と組み合わせて行うデザインの研究の質評価としては参考になる。

[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。

** 泉質の記載については、すべての研究デザインで共通する事項であるため省略した。

ところで、*Nature*や*Science*のような有名科学雑誌に掲載される論文は、よりコンパクトに記載されている。しかし、チェックリスト・声明は、むしろ多くの情報を記載することを要求している。研究者は、疫学・臨床研究における論文の形式を再認識すべきかもしれない。

学術雑誌の中には、規定されたページ数を超えると、超過料金が高価なものがある。これが前述のように、論文中に記載しなければいけない事項（内容）を、実際には適正に実施されていながらも、金額のために省略（文字数を減らす）している状況があるとすれば、結果として、論文と学術雑誌の双方に大きなダメージを与えていることになる。ひとつの手段として電子ジャーナル化することも考えられる。いずれにしても、著者自身や学術雑誌のあり方に直結する問題である。

V パブリケーション・バイアス

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こと、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いうことである。

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ムティック・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、研究を実施したが仮説どおりの結果ではなく、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に出てこない状況にあった。

しかし、1999年から*BMJ*³⁰⁾などは、事前の登録を推奨している。2004年の「パキシル®」服用による小児の自殺企図が報告されなかったスキャンダルを契機に、急速に臨床試験の登録・公開制度が重要視されるようになった。

2004年には、先のICMJEが、臨床試験の登録（clinical trial registry: CTR）を論文投稿の条件とした³¹⁾。WHOも2005年にジュネーブ会議を開催し、ポリシーと基準を定めた³²⁾。

日本においても、パブリケーション・バイアスを回避するために、「UMIN-臨床試験登録」（UMIN-CTR）³³⁾が開始され、年度ごとに登録件数も増加している。これはパブリケーション・バイアスの減少、倫理、臨床試験参加の推進の3つが目的であるが、SRのレビューアが、包括的

Table 4 「お勧め度」のグレード（文献 35, 36 より一部改変*作表）

A	しなさい (Do it)	相応の知識のある人々であれば、ほぼ全員が「実施する」と判断する方法。
B	した方がよい (Probably do it)	相応の知識のある人々の大半が「実施する」と判断するが、その一方で「実施しない」の判断を下す人々が一部存在する方法。
C	しない方がよい (Probably don't do it)	相応の知識のある人々の大半が「実施しない」と判断するが、その一方で「実施する」の判断を下す人々が一部存在する方法。
D	するな (Don't do it)	相応の知識のある人々のほぼ全員が「実施しない」と判断する方法。

*「アルファベット」の記載と「方法」という言葉の加筆。

網羅的に漏れなく論文(研究)を収集する助けにもなる。温泉研究においても、RCTに限らず臨床研究においては、UMINへの事前の登録を推奨することが必要である。

SRにおいても情報バイアスがある。代表的なものは、レビュー・バイアスで、Verhagenら³⁰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。研究者は、著名な雑誌を特別視することの潜在性は否定できないところであり、レビューにおいてはとくに注意を要する点である。

VI お勧め度のグレードと有害事象

お勧め度は、エビデンスや他の要素を含めた意思決定(decision making)や判断(judgment)のためのグレーディングである。これも複数存在し、それらの作成者間で協議がなされ、2004年にGRADEプロジェクト(<http://www.gradeworkinggroup.org/>)が形成されている。

お勧め度のグレードをTable 4に示す。温泉研究としては、観察研究として有害事象のより正確な発生状況を把握するとともに、介入研究ではその有効性と有害事象を並列して示す必要があるだろう。

温泉は、「害よりも効果が上回る」という点では有利な場合も多く、独自のお勧めグレーディング・テーブルの作成も可能かもしれない。合わせて、費用-便益、費用-効果、費用-効用の評価も、今後、温泉研究では求められると考えられる。

附記

本研究は、第72回日本温泉気候物理医学会シンポジウム「温泉気候物理医学のエビデンスを構築する」(2007年5月19日、箱根)において、発表題目「エビデンス・グレーディングの流れ

と温泉医学」で報告した内容を一部含んでいる。

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、司書兼サーチャーとしてご協力いただきました首都大学東京図書館情報センター荒川館の眞喜志まりと東京厚生年金病院図書館の山田有希子の両氏に深謝いたします。

参考文献

- 1) 財団法人厚生統計協会: 国民衛生の動向. 2004, p. 12.
- 2) Moher D, Schulz KF, Altman DG, et al.: The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001; 285: 1987-1991.
- 3) 津谷喜一郎, 小島千枝, 中山健夫(訳): 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> 2002: 118-124.
- 4) Moher D, Jones A, Lepage L, et al.: Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA* 2001; 285: 1992-1995.
- 5) Begg C, Cho M, Eastwood S, et al.: Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996; 276: 637-639.
- 6) 津谷喜一郎, EBMにおけるエビデンスの吟味. *Therapeutic Research* 2003; 24: 1415-1422. [<http://www.lifescience.jp/ebm/opinion/200308/>]
- 7) Moher D, Cook DJ, Eastwood S, et al.: Improving the quality of reports of meta-analyses

Table 3 非ランダム化比較試験と観察研究の質を高めるためのチェックリスト・声明と特徴及び温泉研究への応用性

研究デザイン	名称または略称	特徴と温泉研究への応用性*
非ランダム化比較試験 (nRCT)	TREND声明 ²¹⁾	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nRCTが実施しやすい。温泉研究も同様な状況にあるので、参考にすべきチェックリストである。
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 ²²⁾²³⁾	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(例えば、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。観察研究のデザインは、温泉研究においても論文数が多く、それらをレビューする上で参考になる。
観察研究(OS)	STROBE声明 ²⁴⁾²⁵⁾	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。温泉医学の研究がこうした研究デザインで多数実施されていることを考えると、準拠して記載することの意義は大きい。

[注]

* 泉質の記載についてはすべての研究デザインで共通する事項であるため省略した。

日本温泉気候物理医学会は、「温泉の医学領域への応用とその評価」として、アウトカムとして詳細な臨床検査項目や心理テストのリストを示している²⁶⁾。今後の研究の質を高め、エビデンスをより一層積み重ねることを学会のポリシーとするならば、これらを有効活用・発展させるとともに、Table 2-3を参考にした温泉医学独自のチェックリストや声明の作成が必要と考えられる。

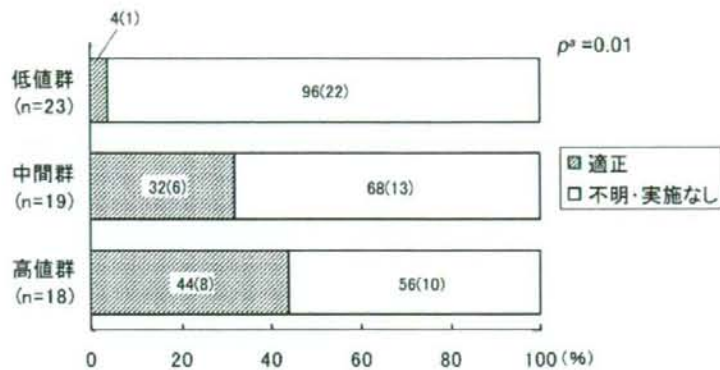
IV 論文の質と方法の質の一致度

Chalmersらは、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²⁷⁾。“Chalmer's eleven point check list”として知られている。

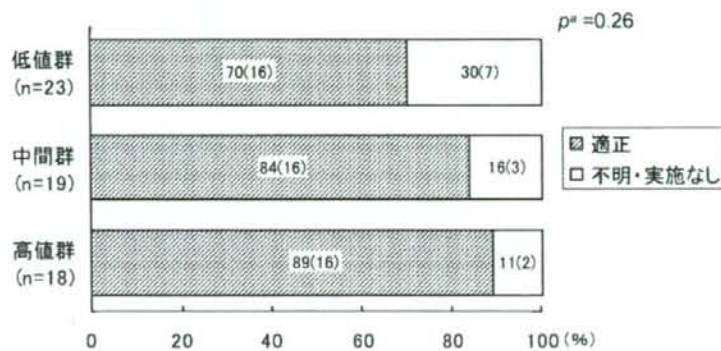
Huwiler-Müntenerら²⁸⁾は、論文全体としては良い報告であっても、方法の質 (concealment: 隠蔽、ブラインド化、intention-to-treat分析: ITT分析)が高いとは限らないことを60編のRCTの分析から報告している (Fig.2)。ブラインド化については、合計得点の高値群・中間群・低値群で有意差はなかった。また、On treatment

分析とITT分析をともに実施した33編の得点の中央値は15点だったが、On treatment 分析だけしか実施していない14編の中央値は14.5点で、両群間に差はなかった (P=0.67: 図表の記載なし)。こうしたことは、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している。

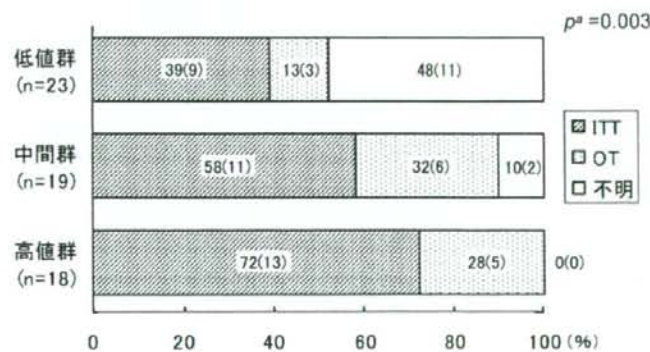
Soaresら²⁹⁾は、悪性腫瘍での生存と患者のQOLを高めるための研究を推進するRadiation Therapy Oncology Groupによって行われたRCT58編を分析した。その結果、事前のサンプルサイズの計算を実施していたのは44編(76%)だったが、論文中にそれを記載していたのは、わずか9編(14%)であり、同様に隠蔽については58編すべてで実施されていたが、記載してあったのは24編(41%)だったことを報告している。掲載されている論文は、必ずしも実施したことをすべて掲載しているとは言えず、臨床研究登録の研究計画書を調べたり、研究者に直接問い合わせることも重要であるとし、合わせ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。



A. 隠蔽 (concealment) の状況



B. ブラインド化の状況



C. ITT分析実施の状況

Fig.2 論文の質得点別に見た各種の分析の質 (Huwiler-Müntener ら²⁸⁾, より著者作図)
 値は% (度数), ^a: χ^2 検定, OT: On treatment 分析

論文の質は, 25点満点中, 低値群 3.5-10.5点, 中間群 11.5-16点, 高値群 16.5-22点である

- of randomized controlled trials : the QUOROM statement. *Lancet* 1999 ; 354 : 1896-1900.
- 8) 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.
- 9) Campbell MK, Elbourne DR, Altman DG, et al. : CONSORT statement : extension to cluster randomized trials. *BMJ* 2004 ; 328 : 702-708.
- 10) Piaggio G, Elbourne DR, Altman DG, et al. : Reporting of noninferiority and equivalence randomized trials: An extension of the CONSORT Statement. *JAMA* 2006 ; 295 : 1152-1160.
- 11) Ioannidis JPA, Evans SJW, Gotzsche PC, et al. : Better reporting of harms in randomized trials: An extension of the CONSORT Statement. *Ann Intern Med* 2004 ; 141 : 781-788.
- 12) Boutron I, Moher D, Tugwell P, et al. : A checklist to evaluate a report of a nonpharmacological trial (CLEAR NPT) was developed using consensus. *J Clin Epidemiol* 2005 ; 58 : 1233-1240.
- 13) MacPherson H, White A, Cummings M, et al. : Standards for reporting interventions in controlled trials of acupuncture: the STRICTA recommendations. *Complement Ther Med* 2001 ; 9 : 246-249.
- 14) 津嘉山洋, 山下仁 : 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 2002; 52 : 582-586.
- 15) Physiotherapy Evidence Database: Access, <http://www.pedro.fhs.usyd.edu.au/index.html>.
- 16) 下坂充 : EBPTにおけるデータベースの活用方法. 理学療法 2001 ; 18 : 1036-1042.
- 17) 上岡洋晴, 黒柳律雄, 小松泰喜ら : 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 2006 ; 69 : 155-166.
- 18) Gillespie LD, Gillespie WJ, et al. : Interventions for preventing falls in elderly people (Cochrane Review). The Cochrane Library, Issue 3, 2004. Chichester, UK: John Wiley & Sons, Ltd.
- 19) Latham N, Anderson C, Bennett D, et al. : Progressive resistance strength training for physical disability in older people (Cochrane Review). The Cochrane Library, Issue 3, 2004. Chichester, UK: John Wiley & Sons, Ltd.
- 20) 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 2007 ; 44 : 403-414.
- 21) Jarlais DC, Lyles C, Crepaz N, et al. : Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions. *Am J Public Health* 2004 ; 94 : 361-366.
- 22) Stroup DF, Berlin JA, Morton SC, et al. : Meta-analysis of observational studies in epidemiology. *JAMA* 2000 ; 283 : 2008-2012.
- 23) 中山健夫 (訳), 津谷喜一郎 (監修) : 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> 2002 : 52-58.
- 24) Elm E, Altman DG, Egger M, et al. : The strengthening the reporting of observational studies in epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Ann Intern Med* 2007 ; 147 : 573-577.
- 25) 上岡洋晴 (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. 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, 2008. (印刷中)
- 26) 鏡森定信, 飯嶋正広, 上馬場和夫 : 温泉の医学領域への応用とその評価. 日本温泉気候物理医学会編集 : 新温泉医学. pp.12-21, JTB 印刷株式会社, 東京, 2004.
- 27) Chalmers TC, Smith HJ, Blackburn B, et al. : A method for assessing the quality of a randomized control trial. *Control Clin Trials* 1981 ; 2 : 31-49.
- 28) Huwiler-Müntener K, Juni P, Junker C, et al. :

Only three papers were found from the search results. Siedliecki (2006) described locomotrium pain and depression in subjects relieved by listening to music (8 pts), Fitzsimmons (2001) described depression in the elderly relieved by moving around on the "Easy Rider Wheelchair Bike" (8 pts), and Parker et al (2001) reported the effects of leisure therapy and occupational therapy on the mood, leisure participation, and independence in activities of daily living of stroke patients (11 pts).

Although only three RCTs were found, numerous non-RCTs and observational studies on the psychosomatic effects of leisure activity and recreation existed both in Japan and abroad. Those studies have suggested some beneficial effects of leisure activity and recreation such as improvement of QOL, health-promoting, educational and therapeutic effects. Additional studies with high evidence-graded study design would be necessary to validate the significant suggestions from the previous studies.

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てきた。

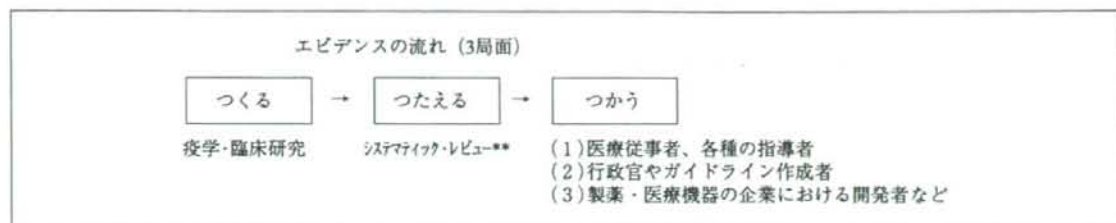
エビデンスの構築と整理のために、研究デザインによってエビデンス・グレーディング(格付け)¹⁾がなされるようになった(表1)。ここでのエビデンス・グレーディングは、ヒトゲノムなどの基

表1 エビデンス・グレーディング(格付け)¹⁾

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場合の例(架空)

- I IIに基づく複数の研究結果を選択的・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。
(縦断的研究)
- 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。
(群間の比較研究、横断的研究)
- 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



* 変更部分は構図の変更。

** より質の高い研究(ランダム化比較試験など)を選択的・網羅的に収集・統合し、批判的吟味を行い、結果を解釈して、公表すること。

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 より一部変更* 作図)

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、生物統計学などの研究デザインが対象となる。仮にレジャー活動やレクリエーションが、心身に及ぼす効果を探る研究の場合には、研究デザインの段階において、これを見越して実施することが、エビデンスを語る上で重要である。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、次いで「ランダム化比較試験(RCT)」、そして最下位が「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」となっている。例えば、「Iの研究とV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」として採択されることになる。

最上位とされるRCTのSRは、とくに臨床医学に関連する研究としては「コクラン共同計画*」²⁾、福祉・教育・刑事司法などの研究領域や学際領域としては「キャンベル共同計画**」³⁾がよく知られており、世界中のRCTを選択的・網羅的に収集し、SRに基づいて介入による効果を評価している。

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⁴⁾。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と考えられる成果を現場に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(とくにRCTデザイン)」と「良質な研究結果を正確にまとめること(SR)」に相当する。

ところで、総説論文においては、従来からの記述的レビュー(narrative review)とシステマティック・レビュー(systematic review)⁴⁾がある。前者は、論文の収集・採用時において、著者の主観的な選択や筆者の考えを支持する論文を選択的に採用する傾向が強いこと、また質の低い論文まで議論の対象とするなどの問題点があるのに対して、後者は、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

コクラン共同計画*

1992年に発足し、独自にシステマティック・レビューであるコクラン・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」⁵⁾を目標とし、システマティック・レビュー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(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。

キャンベル共同計画**

2000年に発足し、コクラン共同計画と姉妹機関として密接な関係を有している。ランダム化比較試験と非ランダム化比較試験を明確に分け、システマティック・レビューによって、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、更新をしていく世界的な評価プロジェクトである。

結果を統合して、批判的吟味とともに一定の結論づけができるところに大きな特徴がある。

これまで「レジャー活動」と「レクリエーション」に関する前者の研究は数多くなされているが、後者のRCTのSRは行われていない。「レジャー活動」と「レクリエーション」による疾病の治療や健康増進の効果、あるいは青少年への教育効果など、効果のエビデンスを述べるにはSRが必要である。ここで扱う「レジャー活動」と「レクリエーション」の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。ただし、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まれない。

本研究は、「レジャー活動」と「レクリエーション」の心身に及ぼす効果についてのランダム化比較試験(RCT)のシステマティック・レビュー(SR)を行うことにより、1) 効果を明らかにすることと、2) 研究の質を評価することを目的とした。

なお、本論では、「evidence (エビデンス)」の用語の定義を「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」を考究するものである。

2. 方法

適格基準を表2に示す。英文キーワード検索として、“leisure activity” and “randomized controlled trial”と“recreation” and “randomized controlled trial”、和文キーワード検索としては、「レジャー活動」と「ランダム化比較試験」と「レクリエーション活動」と「ランダム化比較試験」として行った。データベースは、OVID (full text)、Web of Science、PubMed、Scopus、医学中央雑誌、JDream IIであった。適格基準は、研究デザイン:RCT、出版言語:無制限、対象・サンプル数:無制限、観察期間:無制限、評価指標:無制限であった。

除外基準は、「ランダム化比較試験ではないこと」、「レジャーやレクリエーションの言葉が論文中にあっても、その定義がなく、身体活動や運動、リハビリテーションが主な介入方法であること」、「レジャー活動やレクリエーションとは無関係な内容であること」であった(表3)。検索は、2006年6月から9月の期間に実施した。

RCT研究の質を評価するために、「PEDro Scale」⁶⁾および上岡ら⁷⁾と高橋ら⁸⁾の先行研究に基づいて、17項目からなる評価指標を作成した。各項目について、「実施していれば1点」、「実施していない、または、記述がなければ0点」の17点満点

表2 適格基準

1. 研究デザイン	ランダム化比較試験(RCT)
2. キーワード	<英文> recreation and randomized controlled trial leisure activity and randomized controlled trial <和文> レクリエーション and ランダム化比較試験 レジャー and ランダム化比較試験
3. 出版言語	無制限
4. 対象・サンプル数	無制限
5. 観察期間	無制限
6. 評価指標	無制限

表3 除外基準

1. 研究デザインが、ランダム化比較試験ではないこと。
2. 「レジャー活動」「レクリエーション」の言葉は文中にあっても、その定義がなく、身体活動や特定の運動・リハビリテーションが主な介入方法であること。
3. 「レジャー活動」「レクリエーション」の内容ではない研究であること。

表4 英文・和文によるキーワード検索結果

使用データベース	検索年月	ヒット数	最終の論文数
OVID (full text)	2006年6月	196	2
Web of Science	2006年6月	18	0
PubMed	2006年6月	37	1*
Scopus	2006年9月	13	1
医学中央雑誌	2006年9月	119	0
JDream II	2006年8月	8	0

* OVID と重複。

で評価した。

3. 結果

「レジャー活動」と「レクリエーション」のキーワードを合わせて、英語検索では「Ovid (full text)」が196件、「Web of Science」が18件、「PubMed」が37件、「Scopus」が13件、日本語検索では「医学中央雑誌」が119件、「JDream II」内の「JMedPlus」が8件ヒットしたが、実際に該当する論文は3編だけであった(表4)。

Siedliecki⁹⁾は、運動器の良性かつ慢性的な疼痛を有する患者60名を対象として、自分で好きな音楽を聴く群(PM)、実験者がリラクゼーション

になると想定した音楽を聴く群(SM)、対照群(C)に割り付けし、1日1時間、7日間連続で聴かせた結果、Cと比較してPMとSMが有意な疼痛と抑うつ軽減、活気・生活機能の向上があったが、PMとSMの間には有意差はなかったことを報告している(表5)。

Fitzsimmons¹⁰⁾は、介護福祉施設に入所している抑うつ傾向にある高齢者40名を対象として、車椅子連結自転車を利用して2週間(1日1時間で週5回)レクリエーション(散歩)をする群(recreation: R)と対照群(control: C)に割り付けた結果、Rで抑うつの程度が有意に軽減したことを報告している(表6)。

表5 エビデンス・テーブル(1)

著者	Siedliecki SL
発表年	2006
タイトル	音楽が活力、疼痛、抑うつ、生活機能障害に及ぼす効果
目的	慢性疼痛のある良性の運動器疾患患者(CNMP)を対象として、自分で選択した曲の音楽鑑賞群と実験者が準備した曲の鑑賞群、対照群のいずれが、介入後の活力、疼痛、抑うつ、生活機能障害に効果があるかを明らかにすることを目的とした。
方法	CNMPをもつ21~65歳のアフリカ系アメリカ人と白人患者男女60名を、緊張緩和・精神安定または気分高揚を目的として自分で好みの曲を選択する群(PM:N=22)、実験者が選んだリラクゼーション音楽を提供する群(SM:N=18)、対照群の(N=20)の3群にランダム割付を行い、PM群とSM群については1日1時間、7日間連続で音楽を聴かせた。各群について介入前後に活力、疼痛、抑うつ、生活機能障害の指標を測定した。
評価尺度 (アウトカム)	疼痛:McGill Pain Questionnaire short form 抑うつ:Center for Epidemiology Studies Depression scale 生活機能障害:Pain Disability Index 活力:Power as Knowing Participation in Change Tool
結果	音楽を用いた2群(PMとSM)は、対照群と比較して、活力(p<0.05)が有意に向上、疼痛(p<0.01)と抑うつ(p<0.01)が有意に低下、生活機能(p<0.05)の有意な回復があった。しかし、すべての項目において、PM群とSM群間では有意差はみられなかった。
結論	看護師(関係者)は、CNMPの患者に対して、痛み、抑うつ、生活機能障害の軽減や活力の増進を目的として音楽(曲)の用い方を指導することが望まれる。

表6 エビデンス・テーブル(2)

著者	Fitzsimmons S
発表年	2001
タイトル	車椅子連結自転車を利用したレクリエーション療が高齢者の抑うつ状態に及ぼす効果
目的	米国ニューヨーク州の福祉施設に長期入所している抑うつ傾向のある高齢者を対象として、車椅子連結自転車を利用した2週間のレクリエーション療が高齢者の抑うつ状態に及ぼす効果を検証する。
方法	対象：うつ診断や症状を呈する入所者の中で、GDS簡易版が4点以上であった40名(男性72%、平均年齢80.5歳、平均入所期間29.1年)をランダムに介入群と対照群に割付した。 介入：2週間(1時間×週5日)にわたり、車椅子と自転車を前後に連結した車椅子連結自転車を用い、スタッフが自転車を操縦して屋内外を散策するレクリエーション(乗車時間は1人15分、前後にグループで談話)を実施した。
評価尺度 (アウトカム)	short-form Geriatric Depression Scale (GDS簡易版) 高齢者の抑うつ尺度で、身体、感情、意欲、自尊心、希望に関する15項目(はいいいえ)からなり、5点以上で抑うつ状態と評価される。
結果	介入後のGDS点数は、介入群(19名、脱落1名)で有意に低下($p<0.0001$)し、対照群(20名)ではわずかに上昇しており、介入による抑うつ状態の改善が示唆された。 下位分析により、認知症(介入群の内11名)の者及び抑うつ状態が軽度(診断や服薬を受けていない)の者ほど改善傾向が高く、薬剤による治療の前に、心理・社会的介入を活用することの有効性が示唆された。 女性よりも男性の方が、介入に興味を示す傾向がみられた。
結論	車椅子連結自転車を利用したレクリエーション療法は、長期入所高齢者の抑うつ状態を改善することが示唆された。

表7 エビデンス・テーブル(3)

著者	Parker CJ, et al.
発表年	2001
タイトル	脳卒中後のレジャー療法と一般的な作業療法の多施設ランダム化比較試験
目的	脳卒中後の患者に対するレジャー療法と作業療法による介入で、気分、レジャー参加、日常生活動作(ADL)を退院後、12ヶ月後まで追跡して明らかにすることを目的とした。
方法	対象：5つの病院で466名。レジャー(L)群：163名(平均年齢72歳)、ADL(ADL)群：156名(平均年齢71歳)、対照(C)群：157名(平均年齢72歳)にランダム割付した。 介入(作業療法士による)：ADL群は退院後最長6ヶ月まで、とくにADLの向上を目的として10回以上(1回30分以上)の家庭用プログラムを実施させた。L群は退院後最長6ヶ月まで、レジャー活動への参加とそれが促進されるような動作要素の訓練を実施させた。C群には何も介入は行わなかった。ただし、すべての者は外来診察のように、地域でのリハビリテーションなどは自由に受けられるものとした。
評価尺度 (アウトカム)	気分：General Health Questionnaire (GHQ) レジャー活動(量と頻度)：Nottingham Leisure Questionnaire (NLQ) ADL：Nottingham Extended ADL Scale (NEADL) など
結果	12ヶ月後まで追跡できたのは、L群：74%(113名)、ADL群：68%(106名)、C群：71%(112名)であった。 GHQ(気分)は6ヶ月後、12ヶ月後において、C群と比較して、L群・ADL群はともに有意差はなかった。NLQ(レジャー活動)も6ヶ月、後12ヶ月後において、C群と比較して、L群・ADL群はともに有意差はなかった。NEADL(ADL)も6ヶ月後、12ヶ月後において、C群と比較して、L群・ADL群はともに有意差はなかった。
結論	先行研究に反して、退院後の脳卒中患者に対する追加的な作業療法(ADLとレジャー)は、気分・レジャー活動・ADLに有意な効果をもたらさなかった。

表8 研究の質についての評価（文献6-8による評価項目を用いた）

大項目	小項目	Siedliecki (2006)	Fitzsimmons (2001)	Parker et al. (2001)
対象者選定 と ランダム化	1)対象選定(採用と除外)の基準は明記されているか?	1	1	1
	2)レジャーあるいはレクリエーションの定義づけがなされているか?	0	0	0
	3)ランダムなグループ分けがされたか?	1	1	1
	4)群の割付について隠蔽がなされたか?	0	0	1
	5)両群はベースラインで同等だったか?	1	1	1
盲検化	6)対象者は盲検化されたか?	0	0	0
	7)介入者は盲検化されたか?	0	0	0
	8)評価者は盲検化されたか?	0	1	1
介入・測定	9)介入内容(種類、頻度、時間、期間、場所、強度など)の詳細は明記されているか?	1	1	0
	10)サンプル数は十分か? (ベースラインで各群50以上、 またはパワー分析に基づき決定したか?)	0	1	1
	11)観察期間は十分か?(3ヶ月以上)	1	0	1
	12)測定・評価方法は明記されているか?	1	1	1
分析・結果	13)主な指標においてベースライン時の対象者の85%以上の測定がなされたか?	1	1	0
	14)主な指標においてIntention-to-treat分析がなされたか?	0	0	1
	15)主な指標において統計学的群間比較がなされたか?	1	0	1
	16)主な指標において点推定と信頼区間の両方を示しているか?	0	0	1
	17)有害事象に関する記述があるか?	0	0	0
合計得点	---	8	8	11

Parkerら¹¹⁾は、脳卒中患者(466名)に対して、退院後に6ヶ月間の一般的な自宅での作業療法を行う群(1回30分以上で10回以上)と、レジャー活動を行う群(レジャー活動に必要な動作要素も含む)、何もしない対照群の3つに割り付けをして、気分や日常生活動作の程度、レジャー活動の参加の程度を介入終了後6ヶ月後まで追跡して調べた結果、群間に有意差がなかったことを報告している(表7)。作業療法士は、自己の治療について過大評価をせずに、効果が生じるかどうか

を常に観察する必要性があることも合わせて示している。

表8は、研究の質の評価である。3研究に共通しているのは、「レジャー活動とレクリエーションの定義がなされていない」、「有害事象の記述がない」という事項であった。

4. 考察

システマティック・レビューの基本的な手順に従って、選択的・網羅的な論文収集を行った。著

名で大規模な雑誌を包括する6つのデータベースを使用したことから、該当論文の漏れが生じていることは考えにくい。データベースにおいて、ヒットした論文数は多いものの、抄録 (abstract) をすべて読み、除外基準に従った結果、3編だけとなった。シソーラスの「下位語も含めた検索を行う」機能を用いてキーワードに基づき検索を行ったが、例えば、「レジャー」のシソーラスは、下位語に指定されている「休日」、「リラクゼーション」、「休息」、「趣味」、「スポーツ」などの用語が含まれている文献もヒットしてしまうことが、実際に適格基準に適した3編の論文とヒットした論文数の差異として生じた原因のひとつと考えられる。また、論文中にこれらの用語が出てると、例えば、考察の中で先行研究のRCTを引用して記述していると、RCTという用語があるのでデータベースはこれを取り込んでしまうことも数が多くなった理由として考えられた。

研究が少なかったことは、RCTデザインで実施することが困難だということだけではなく、「レクリエーションとレジャー活動」という用語について学術的なコンセンサス（とくに医療・保健・福祉や関連する学際領域）が国際的に得られていないため、「運動」や「リハビリテーション」などの別の用語にマスクされている可能性が高いためと考えられる。国内外を問わず、「レジャー活動」と「レクリエーション」による人に対する効果のエビデンスを述べる上では、この用語が運動・身体活動やリハビリテーションなどの用語と完全区別され、独立した介入方法となりうる学際的なコンセンサスが必要だろう。

現時点においてRCTデザインの研究は少ないものの、非ランダム化比較試験あるいは症例対照研究やコホート研究などの観察研究は数多く実施されているはずであり、治療効果や健康増進効果、教育効果のエビデンスを明らかにするためには、こうした研究を含むSRも必要である。本研究は、メタ分析による統合をも計画したが、それには至らなかった。システムティック・レビューは、定期的あるいは新たに更新していくことが必要であるため、今後の課題となった。

研究の質をみると、17点中、11点が1編、8点が2編であり、中庸と考えられる研究の質であ

った。しかし、「レクリエーション」の定義づけがすべての研究でなされていないことは、前述のように「レクリエーション」の効果と明確にできない限界を含んでいた。

アウトカムにおいて2つの研究^{8) 9)}に共通している点は、抑うつ症状の改善が含まれていたことだったが、「虚弱高齢者が楽しめるように工夫された特殊な自転車に乗せてもらっての散歩」と「音楽を聴くこと」がレクリエーションと位置づけられ、異なる介入方法であったために、メタ・アナリシスによる統合を行うことは不可能であった。レジャー治療と作業治療の効果については、唯一の研究のため統合を行うことはできない。こうしたことから、それぞれの効果を個別に示しているにすぎないSRとなっており、この点が本研究の限界ともなっている。さらに検索が2006年6月から9月と3ヶ月間を要しており、タイムラグが生じていることも本研究の限界である。

5. まとめ

「レジャー活動」と「レクリエーション」が心に及ぼす影響について明らかにしたランダム化比較試験は、看護の欧文雑誌2編と臨床リハビリテーションの欧文雑誌1編に掲載されていた。前者2編では、高齢者における運動器の疼痛軽減とうつの改善にレクリエーションは効果があったとし、後者1編は脳卒中患者において、レクリエーション活動参加と作業療法との間に、日常生活動作の回復過程の効果は有意な差がなかったとする報告であった。エビデンス・グレーディングの高い研究デザイン2編から、レクリエーションは、運動器の疼痛の軽減とうつ症状を軽減させる可能性があることが示唆された。研究の質は、それぞれ17点満点で、8点、8点、11点であった。

「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研究デザインは、現段階においては少ないが、非ランダム化比較試験や観察研究は多数報告されている。これまでの「レジャー活動」と「レクリエーション」による教育効果、治療効果、健康増進効果、QOLの向上などの意義ある報告を、今後は、よりエビデンス・グレーディングの高い研究デザインによって効果を実証していく必要性が示された。

附記

本研究は、日本レジャー・レクリエーション学会第36回大会で発表した内容を一部改変したものである。

平成18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。

参考文献

- 1) 財団法人厚生統計協会：国民衛生の動向, p.12, 2004.
- 2) コクラン共同計画：アクセス,
<http://www.cochrane.org/>
- 3) キャンベル共同計画：アクセス,
<http://www.campbell.gse.upenn.edu/>
- 4) 津谷喜一郎：EBMにおけるエビデンスの吟味, *Therapeutic Research*, 24 (8) :1415-1422, 2003.
- 5) 津谷喜一郎：エビデンスを調べる-systematic review の現状-, *臨床薬理*, 34 (4) :210-216, 2003.
- 6) Physiotherapy Evidence Database: Access, <http://www.pedro.fhs.usyd.edu.au/index.html>
- 7) 上岡洋晴, 黒柳律雄, 小松泰喜他：温泉の治療と健康増進効果に関するシステマティック・レビュー, *日本温泉気候物理医学会誌*, 69:155-166, 2006.
- 8) 高橋美絵, 上岡洋晴, 津谷喜一郎他：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 (4) :403-414, 2007.
- 9) Siedliecki SL: Effect of music on power, pain, depression and disability, *J Advanced Nursing*, 54: 553-562, 2006.
- 10) Fitzsimmons S: Easy rider wheelchair biking: a nursing-recreation therapy clinical trial for the treatment of depression, *J Gerontol Nurs*, 27:14-23, 2001.
- 11) Parker CJ, Drummond AER, Deway ME et al: A multicentre randomized controlled trial of leisure therapy and conventional therapy after stroke, *Clin Rehabil*, 15:42-52, 2001.

(受付：2007年3月12日)
(受理：2007年11月5日)

疫学・臨床研究デザインと司書の役割

東京農業大学地域環境科学部身体教育学研究室
上岡 洋晴

I. エビデンスのグレーディング

1990年代後半から「科学的根拠に基づいた医療 (Evidence Based Medicine: EBM)」や「科学的根拠に基づいた健康政策 (Evidence Based Health Policy: EBHP)」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てきた。

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のグレーディング (格付け)¹⁾ がなされるようになった (表1)。

ここでのエビデンス・グレーディングは、ヒトゲノムなどの基礎医学ではなく、人間を対象とした疫学研究や臨床研究の研究デザインが対象となる。

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の「システマティック・レビュー (Systematic Review: SR)」(メタ・アナリシス (Meta-Analysis) を含む) であり、次いで「ランダム化比較試験 (Randomized Controlled Trial: RCT)」,そして最下位が「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」となっている。例えば、「Iの研究とVI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。

最上位とされるRCTのSRは、とくに臨床医学

に関連する研究としては「コクラン共同計画²⁾」²⁾、保健・福祉・教育・刑事司法などの研究領域や学際領域としては「キャンベル共同計画²⁾」³⁾がよく知られており、世界中のRCTを選択的・網羅的に収集し、SRに基づいて介入による効果を評価している。

これらは、研究者は当然ながら、司書においても、医師やコメディカルからの文献の要望や検索の相談に応じる場合に、理解していることの意義は大きいだろう。

II. エビデンスの流れと総説論文

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⁴⁾。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と考えられる成果を現場に役立てることとなる。他方の「つくる」と「つたえる」は、換言す

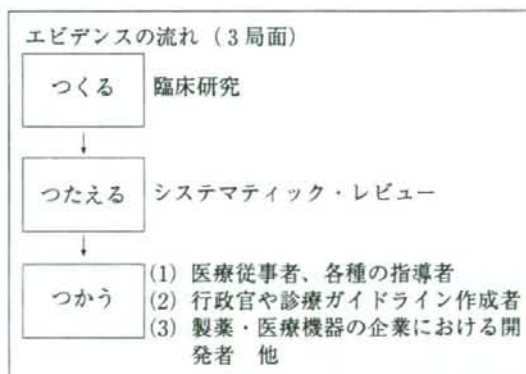


図1 エビデンスの流れとシステマティック・レビュー (津谷、2003より作図)⁴⁾

KAMIOKA Hiroharu, Ph. D
東京農業大学地域環境科学部身体教育学研究室 准教授
〒156-8502 東京都世田谷区桜丘 1-1-1
tel: 03-5477-2587 fax: 03-5477-2587
e-mail: h1kamiok@nodai.ac.jp

(受理日: 2007.11.2)

表1 エビデンス・グレーディング (文献1をもとに脚注を著者加筆)

- 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研究
- III 非ランダム化比較試験による研究
- 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

[注]

〈具体的例〉

- I 複数のランダム化比較試験を用いて、各論文の質の評価を行ったり、メタ・アナリシスによって結果を統合し、効果を明らかにする。
- II 母集団内の対象を介入群と対照群などの複数の群に、抽選などにより、ランダム割り付けし、介入 (治療など) の効果を比較すること。
ヒトであることに由来する様々なバイアスを低減・調整できる最良の方法であるが、倫理的な問題も生じやすく、実施するのが難しいことも多い。
- III 母集団内の対照を抽選などではなく、一定の基準に従って、意図的に複数の群に割り付けし、介入の効果を比較すること。
- IV コホート研究はある母集団を経年(時的)に追跡し、曝露のある群とない群の比較(率比・差など)を行うこと。
症例対照研究は、ある時点において、ある症例を有する者(患者)とそうでない健康者の比較を行うこと。例えば、50名の変形性膝関節症者と50名の同性・同年齢の非疾患者のBMIを比較するというような研究である。
- V 1症例あるいは複数の症例の報告。
- VI 先行研究などの報告に基づかない意見。

れば「研究すること (とくにRCT デザイン)」と「良質な研究結果を正確にまとめること (SR)」に相当する。

ところで、総説論文においては、従来からの記述的レビュー (narrative review) とSRがある。前者は、論文の収集・採用時において、著者の主観的な選択や筆者の考えを支持する論文を選択的に採用する傾向が強いこと、また質の低い論文まで議論の対象とするなどの問題点があるのに対して、後者は、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。論文の網羅的・選択的収集については、あるデータベースを用いて、適格基準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は最終的に全ての研究の結果を統合して、批判的吟味とともに一定の結論づけができるところに大きな特徴がある。SRの数は増え続けている。

上述のように、医療にかかわる現場の司書は、従来からの総説とSRの正しい理解が必要だと考えられる。

III. 論文の質を高めるチェックリスト

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) であるが、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」³⁾ 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。最も著名なのが、22のチェック項目からなる「改訂版 CONSORT (Consolidated Standards of Reporting Trial) 声明」⁴⁾ 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 Intern 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