

よくわかるデータの見方と研究のすすめ方

第16回

系統的レビューとはなにか

独立行政法人国立健康・栄養研究所

佐々木敏 *Sasaki, Satoshi*

東京大学医学教育国際協力研究センター

水嶋春朔 *Mizushima, Shunsaku*

「系統的レビュー」は医学の分野ではそれほど珍しいものではなくなりましたが、栄養学の分野ではまだまだ少ないかもしれません。しかし、食品・食事・栄養素と疾患の関連の有無や程度を評価したり、栄養指導の内容を決めたりするうえで、これらはとても重要な情報を提供してくれます。

今回は、EBN (evidence-based nutrition) の重要な情報源となる「系統的レビュー」について考えてみます。

疾患の診断基準や治療指針のガイドラインなどをみていますと、「系統的レビュー」という言葉が最近よく出てきます。「系統的レビュー」とはなんでしょうか？ また、これらはどのように利用すればよいのか、教えてください。

「レビュー」とは、いままでに行われた研究成果をまとめたものを指し、「総説」とも呼ばれます。いままでに行われた研究成果は論文として発表されています。どの論文を選び、どの論文を選ばないかは総説の結果に影響を及ぼす重要な問題です。また、選ぶときに大切な研究成果をもらしてはなりません。そこで、あらかじめ、論文を検索する規則を定め、もれや偏りが生じないように、系統的に集め、系統的に評価する方法によってつくられる総説があり、「系統的レビュー (systematic review)」と呼ばれています。系統的レビューは、ほかのタイプのレビューよりも信頼度が高いと考えられます。

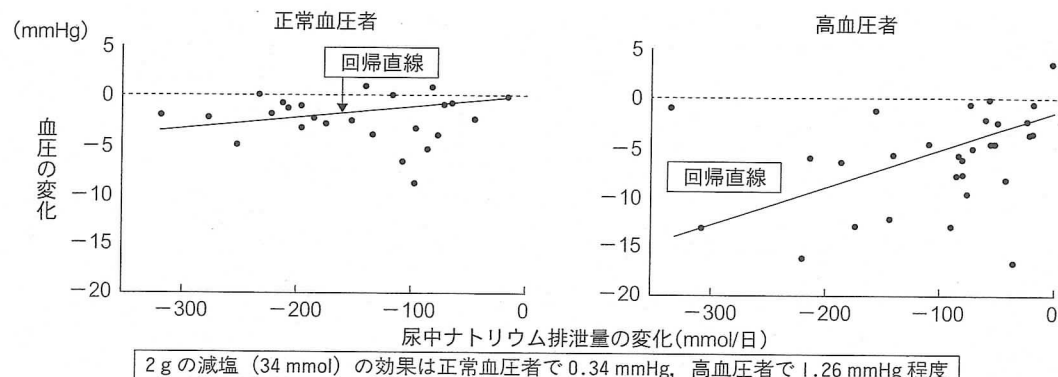
さらに類似の研究結果を数量的に統合して、結果を一つにまとめる作業のことを「メタ分析 (meta-analysis)」と呼びます。これは次々回(18回)に説明しましょう。

各個研究の限界とレビューの発想

ある一つの疑問、たとえば「減塩をすると血圧はどれくらい下がるか」を考えてみましょう。この疑問に答えるために行われた研究は複数あるでしょう。そして、ヒトを対象とした研究では、いくら注意しても、どうしても「偶然」の要素の混入を避けることができず、結果は微妙に変わってしまいます。図1は、ナトリウム摂取量を制限して収縮期血圧の変化を観察した56の研究の結果を図にしたものです¹⁾。これをみると、結果はかなりばらついており、とても大きな降圧効果を示した研究や昇圧効果を示した研究もあります。このような現実をみると、一つひとつの研究(各個研究)の信頼度には限界があることがわかります。

ここに、「一つの研究結果で結論を下すのではなく、たくさんの研究結果を集めてみよう」という発想が生まれます。それが「レビュー」です。しかし、どんな研究でもただ集めればよいというわけではありません。測定方法によって血圧の値が異なるように(3分間座って休むか、経験ある看護師が測定するか、など)、研究結果も研究方法の良し悪しによって異なります。したがって、「〇〇

図1 減塩が収縮期血圧降下に与える効果：56の無作為割付比較臨床試験のまとめ



対象者に減塩をさせて、収縮期血圧の変化を観察した研究の結果をまとめたもの。一つの点の一つの研究を表す。(Midgley et al.: JAMA, 275: 1590-1597, 1996)

以上の質の方法を用いた研究」というように具体的な規則をあらかじめつくっておき、それを満たしたものを選ぶという作業が必要になります。そして、手元にあった論文から選ぶとか、先輩が紹介してくれた論文から選ぶというようなことをしてもなりません。そうではなく、原則的には「世界中の全部の論文」から選ばなくてはなりません。そうでないと、とても重要な研究がもれてしまうことがあるからです。このような手続きを経てつくられたレビューのことを「系統的レビュー」と呼びます。

つぎに、「得られた結果を平均してみよう」という発想が生まれます。平均すれば、もっとも信頼できる代表値が得られるだろうと考えられるからです。このように各個研究から得られた結果を一つのデータとして扱って、それらを数量的に統合して一つの代表的な数値を得る作業のことを「メタ分析」と呼びます。

系統的レビューの実際

系統的レビューで大切なのは「論文の集め方」です。代表的なものはつぎの三つです。

①科学文献データベースを利用して、そこから一定の規則に従って抽出する。

②探したい研究結果が発表されている確率が高い科学雑誌の目次をすべて目で見てもいいです。

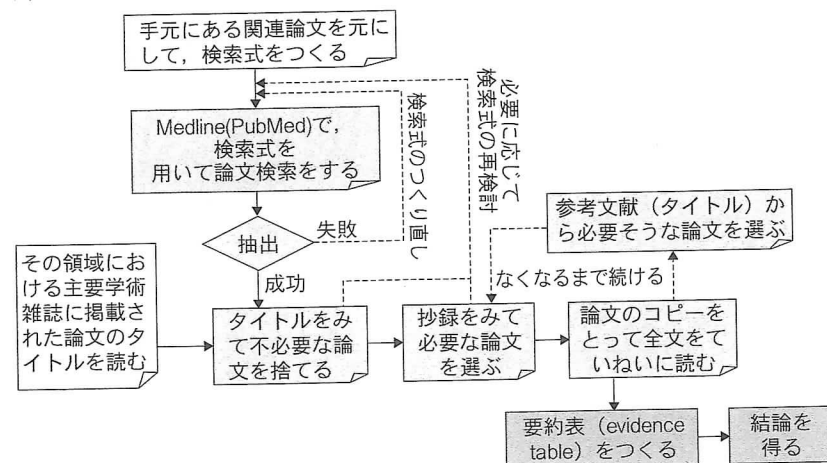
③集められた論文の参考文献リストから目で見てもいいです。

①で使われるデータベースの代表がMEDLINE (Medline)です。MEDLINEは米国国立医学図書館が管理している医学系論文のデータベースで、PubMedという名前でインターネット上に無料で公開され、およそ900万の論文が収録されています (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi>)。そこにアクセスして、キーワードや検索式を入力することによって、そのキーワードや検索式を含む論文を抽出することができます。世界中の主な医学系雑誌はほぼこのデータベースに収録されているため、たいへん効率的に探したい論文を見つけることができます。検索式とは、複数のキーワードを and, or, not, () でつないだ論理式のことです。そして、論文のタイトル(論題)か抄録 (abstract) のなかにその用語が含まれていると、その論文をデータベースが選り出して画面上に表示してくれます。

②は、面倒そうですが、考慮したい方法です。この場合、どの科学雑誌を選ぶかが大切です。たとえば、「栄養」と名前のついた雑誌でなく、注目している疾患に関する雑誌や、公衆衛生や介護など、関連領域を扱った雑誌も忘れてはなりません。

③は、①と②の作業を行い、利用可能な論文が集まった後に、これらの論文の参考文献リストを

図2 Medlineを中心に据えた系統的レビューの基本手順



用いて行います。したがって、これは①と②に対して補足的な意味をもつものと考えておくとよいと思います。

ここまでの作業によって、完全ではありませんが、ほぼ完全に近いレベルで必要な論文を集めることができます。この作業過程を図2に示しておきます。

何について調べるのか

論文を集めるといっても、「何について集めるのか」の疑問が明確でないと集めることはできません(「正しい疑問の作り方」については、2003年2月号の本連載(102巻2号184頁)をご覧ください)。たとえば、「低グリセミック・インデックス(GI)食は糖尿病指導に有効か」という疑問では不十分です。どの指標を用いて有効性を評価するのか(空腹時血糖か、食後血糖か、食事2時間値か、HbA_{1c}の値か、合併症の発症率か)、どのくらいの期間の指導効果をみたいのか(食後2時間か、1週間か、1年間か、生涯か)、どのような状態の患者か(インスリン依存型か非依存型か、若年者か高齢者か、または健常者か)、どのような研究方法を用いた研究か(観察研究か介入研究か、介入研究なら対照群はあるのか、対照群がある場合は無作為割付か)などについて、どの条件で行われた

研究を集めたいのかをあらかじめ決めておかないとはなりません。

表1は、「炭水化物食代謝の変化に関する高GI食と低GI食の無作為割付比較臨床試験」の系統的レビューの結果です²⁾。ここでは食事数時間以内の短期効果ではなく、数週間の試験食における中期効果について検討しています。この結果から、空腹時血糖にはあまり効果がみられなかったが、平均血糖では低下を観察した研究が多かったこと、1980年代の研究では食物繊維摂取量のコントロールが不十分で、観察された効果が低GI食によるものか、高食物繊維食によるものなのかが不明瞭であることなどがわかります。

この表をみると、「低GI食は糖尿病指導に有効」という研究報告があります。「低GI食は糖尿病指導に有効といわれています」というような記述によって与えられる情報との質の差は明らかであると思います。なお、表1のような形式の表を要約表(summary table)と呼びます。エビデンス・テーブル(evidence table)と呼ばれることもあります。

どの情報を集めるのか

MEDLINEでも、探したい研究結果が発表されている確率が高い科学雑誌でも、検索の対象とな

表1 炭水化物代謝の変化に関する高GI食と低GI食の無作為割付比較臨床試験の結果

文献 発表年	著者	対象者	期間	DF摂取量 (%差)*1	GI (%差)*1	空腹時血糖 (%変化)*2	平均血糖 (%変化)*2	HbA _{1c} (%変化)*2	フルクトサミン (%変化)*2
1987	Jenkins et al.	健常者(n=6)	2×2週間	24	-40	NS	-37	—	-7
1987	Jenkins et al.	高脂血症者(n=30)	3×1月間	6	-11	NS	—	—	NS
1988	Jenkins et al.	NIDDM(n=8)	2×2週間	21	-22	-30	—	-7	-7
1988	Calle-Pascual et al.	NIDDMとIDDM(n=24)	2×4週間	—	-6	—	NS	NS	—
1988	Collier et al.	小児IDDM(n=7)	2×6週間	—	-13	NS	—	-27	—
1988	Fontvieille et al.	IDDM(n=8)	2×3週間	—	-14	NS	—	—	-18
1991	Brand et al.	NIDDM(n=16)	2×12週間	0	-13	NS	-14	-11	—
1992	Wolever et al.	NIDDM(n=15)	2×2週間	6	-27	NS	—	—	-3
1992	Wolever et al.	NIDDMと過体重者(n=6)	2×6週間	3	-28	NS	—	—	-8
1992	Fontvieille et al.	NIDDMとIDDM(n=18)	2×5週間	0	-26	-11	-13	NS	-12
1999	Jarvi et al.	NIDDM(n=20)	2×24日間	12	-31	NS	-30	NS	-3
1999	Luscombe et al.	NIDDM(n=21)	3×4週間	0	-32	NS	—	—	NS

*1高GI食に対する低GI食の差。

(文献2より引用)

*2高GI食期間後の変化に対する低GI食期間後の変化の差。

略語：DF=食物繊維(dietary fiber), GI=glycemic index, NIDDM=インスリン非依存性糖尿病, IDDM=インスリン依存性糖尿病, NS=有意差なし。

るのは「論文」です。「論文」にもたくさんの種類がありますが、基本的には「査読制度(peer review)がある科学雑誌に掲載された原著論文(original article)」に限ります。これがもっとも尊重されるべき情報源ですが、しばしば必要に応じて、「査読制度のない科学雑誌に掲載された論文」や「学会発表抄録」「研究報告書」などが用いられることもあります。

読者の立場に立てば、「査読制度がある科学雑誌に掲載された原著論文をていねいに集めて、ていねいに整理したレビューを探し、それがみつからない場合に限って、ほかの方法によってつくられたレビューを読む」ということになります。

まとめ

「疫学」は複数のヒトを扱います。その場合に注意したいことに「集団代表性」と「標準化」があります。調べることができるのは、調べたい人たちのごく一部でしかないのがふつうです。調べることができる人たちが全体の人たちを代表してくれているか否かが「集団代表性」です。一生懸命に測定をしても、測定条件や測定方法が異なれば結果は異なってしまいます。したがって、なにか

を調べるときには、条件や方法をそろえて測定しなくてはなりません。これが「標準化」です。これは、栄養・健康情報を選び、評価する場合にもそのまま当てはまります。つまり、系統的レビューは、「疫学的な考え方を科学情報の整理と理解に応用した例」と考えることができます。

次回は、系統的レビューを読むときの注意点と、系統的レビューに得られた結果をまとめる方法として注目されているメタ分析について説明することにしましょう。

文献

- 1) Midgley, J.P., Matthew, A.G., Greenwood, C.M.T. et al.: Effect of reduced dietary sodium on blood pressure: a meta-analysis of randomized controlled trials. *JAMA*, 275:1590-1597, 1996.
- 2) 佐々木敏: glycemic indexの低い食品は血糖のコントロールに有効か. *EBMジャーナル*, 1: 580-587, 2000.

ご質問募集

当欄は、読者の皆様からのご質問をもとに構成しています。佐々木先生、水嶋先生に質問のある方は、巻末のアンケートはがきか、E-mail(jcn@ishiyaku.co.jp)で編集部宛にお寄せください。質問をとりあげさせていただいた方には、掲載号と図書券をお送りします。