

正しく知ろう! 食の情報



真実を知ることは
豊かで健康的な
暮らしに役立ちます



公益財団法人 日本食肉消費総合センター

〒107-0052 東京都港区赤坂 6-13-16 アジミックビル 5F
ホームページ <http://www.jmi.or.jp>

ご相談・お問い合わせ

e-mail consumer@jmi.or.jp
FAX 03-3584-6865



資料請求 info@jmi.or.jp

畜産情報ネットワーク <http://www.lin.gr.jp>

平成30年度 国産畜産物安心確保等支援事業

後援/aic独立行政法人 農畜産業振興機構

制作/株式会社 エディターハウス

2●栄養疫学研究論文の読み方<メタ・アナリシスの功と罪>

功も罪も含めその論文を読むか
読まないか その論文をどう読むかは
読者側の問題です

東京大学大学院医学系研究科 社会予防疫学分野教授 佐々木 敏



1つの論文だけではわからなくても、多数の論文を集めることによって結論をあぶり出すのがメタ・アナリシスです。「エビデンスに基づく医療」への道をひらくメタ・アナリシスには数多くの功がありますが、さまざまな偏りが生じるという罪の一面向があることも事実と、おっしゃる佐々木敏先生に、メタ・アナリシスの歴史やケーススタディとともに、疫学の研究論文の読み方について、お話をうかがいました。

研究を複数個集めて平均値をとるのがメタ・アナリシス

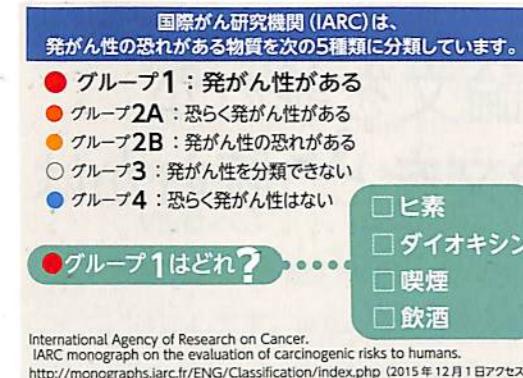
まず初めに考えたいのは、「がん予防のために赤身肉や加工肉は避けるべきか?」という疑問です。2015年10月、WHO(世界保健機関)の下部機関であるIARC(国際がん研究機関)が、「赤身肉とその加工肉には発がん性がある」という論文を発表し、世の中が騒然とした出来事がありました。その結果、市場では赤身肉やその加工品が避けられる事態が引き起こされました。

疫学研究に携わってきた者として、疫学の役割とは何か、研究結果をどのように読むべきかをお話ししなければならないだろうと考えました。IARCが発表した「発がん性の分類」は、グループ1からグループ4まであり、グループ2が

AとBに分かれているので、実際には全部で5グループです。グループ1は「発がん性がある」。グループ2Aは「恐らく発がん性がある」。グループ2Bは「発がん性の恐れがある」。グループ3は「発がん性を分類できない」。グループ4は「恐らく発がん性はない」となっています。

図表1は、ヒ素、ダイオキシン、喫煙、飲酒を列挙し、「グループ1(発がん性がある)はどれ?」と問うています。正解は、全部です。一般的に世の中ではヒ素とダイオキシンが怖いと考えられています。喫煙は、吸わない人が怖いといいます。飲酒になると、お酒を飲む人の評価はかなり甘くなります。これこそが、人を対象とする医学の特徴です。

図表1 國際がん研究機関(IARC)の発がん性の分類



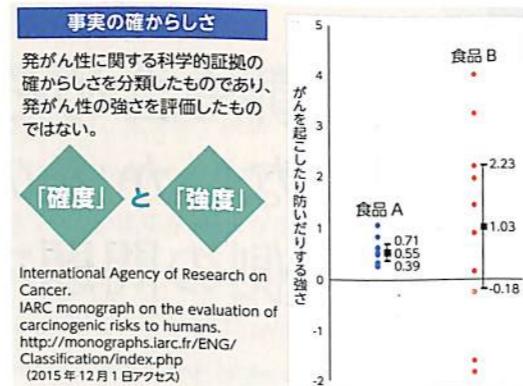
それで、IARCのホームページに入り、発がん性があるとは一体どういう定義か調べてみました。すると、次のように書いてありました。「発がん性の恐れがある5分類について発がん性に関する科学的証拠の確からしさを分類したものであり、発がん性の強さを評価したものではない」と。おわかりでしょう。すなわち、発がん性の確からしさですから、食品Aの発がん性は確かです。食品Bの発がん性は確かではありません。なぜなら有意になっていないからです。けれども発がん性の強い、弱いで見ると、食品Bのほうが食品Aよりも発がん性は強い。

複数の要因を相対的に比較し判断することが重要

図表3は赤身肉・鶏肉と大腸がんの関連をまとめた図です。横軸が赤身肉の1人当たりの摂取量です。といっても、ざっくりとした調査にすぎません。私たち人間が習慣的に食べている赤身肉の量を調べるわけですから、丁寧に測ることはできません。従って相当の誤差があります。特に系統誤差よりも偶然誤差に大きな誤差が起こります。一方、縦軸が発症確率です。

この図の上半分は、大腸がんの発症確率を、

図表2 「確度」と「強度」



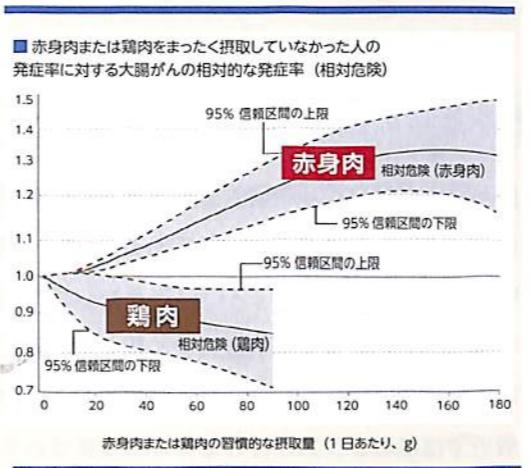
恐らく世間の方はBを怖いという。けれどもIARCはAで決めている。ここが違うんだろうと考えました(図表2)。

すなわちここで重要なのが、1つの研究で結論を下してはならないという疫学の原則で、これこそがメカニズムの原則と大きく違うところです。メカニズムは理由の説明ができる、それが理屈として通っていればそれで証明終了です。けれども、疫学は1つ目の結果で結論を出すことをためらい、避けようします。そして、質の高い研究を複数個集めて、その平均値をとります。これがメタ・アナリシスです。

赤身肉をほとんど食べないと、1日当たり10～20g食べている人の比較です。ほとんど食べない人に比べて、それぞれのグラム数の赤身肉を習慣的に食べている人から発症する大腸がんのリスクを比率として出したもので、相対危険と呼ばれるものです。

すると、発症確率はおよそ1.3倍に上がりました。大腸がんの発症確率が約30%上がるということです。そして、95%信頼区間が明らか

図表3 赤身肉・鶏肉と大腸がんの関連



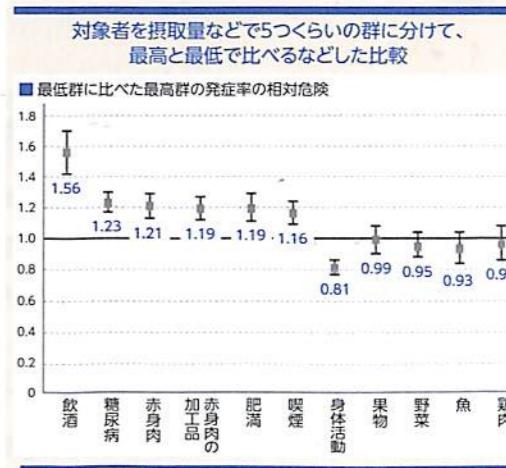
かに1.0を上回っているので、発症確率は有意に上昇し、その大きさは0.3であることがわかります。ただ、120gくらいから後は頭打ちになります、ほとんど上がっていないことがわかります。

一方で、図の下半分にあるように、ホワイトミートの代表である鶏肉の場合は上がりません。この研究では、むしろやや下がるような結果が出ています。ただ、別の研究者が行ったメタ・アナリシスの結果では、鶏肉での発症確率は特に下がっていませんでした。すなわち、メタ・アナリシスも揺れるということなのです。

このデータからいえることは、少なくとも赤身肉は有意に上がり、鶏肉は有意に下がる。けれども、鶏肉の優位の下がり方はわずかであるということです。

大腸がんについては、ほかのリスクについても調べています。これが、人間を扱う私ども医学者の特徴です。すなわち「人は肉だけ食べて生きるにあらず」で、いろいろなものを食べており、いろいろな悪さをしているからです。図表4はいろいろな「悪さ」を並べ、大腸がんを結果として調べた疫学研究です。いろいろな疫

図表4 食習慣などと大腸がんの関連



学研究を集め、そのリスクと相対危険を出してみたものです。

大腸がんにかかりたい人はいないと思います。かかりたくなければ、この中で一体どれに一番注意を払ったらいいと思いますか？ 答えはお酒を飲まないことです。それ以外にありません。その次に糖尿病にかかるないこと。ほぼ同じ強さで、赤身肉を避けること。同時にその加工品を避けること。そして、太らないこととたばこを吸わないことです。

その一方で、大腸がんの発症を予防する重要な因子がありました。それは身体活動、運動をすることです。すなわち、運動することによって避けられる大腸がんと、赤身肉を食べることによって発生する大腸がんはほぼ同じであるということがいえます。

重要なことは、われわれ予防医学者は一つの数字を見るのではなく、複数の要因を相対的に比較し、どちらを優先的に避けるか導入するかを決めるのです。すなわち、ここでは赤身肉がどうという言い方ではなく、赤身肉はほかの危険因子や予防因子に比べてどうかという読み方をしていただきたいのです。

メカニズムを解くのではなく量を決めることが疫学の役割

日本では、これまで疫学はあまり盛んではありませんでした。メタ・アナリシスも、私が国立がんセンターでかかわった研究がありますが、それ以外にそれほど多くはありません。

疫学は医学の中で非常に大きな部分を占めています。薬剤の開発では、理論的開発や動物実験の後、患者さんを使った有効性試験が行われます。

臨床試験ですが、私たちはそれを治験と呼んでいます。臨床試験を経て、実際に薬剤がマーケットに出ます。すなわち、疫学研究こそが薬が世に出る前の大いなる門門になるのです。

ところが、食べ物に関しては臨床試験を経ることなく、「メカニズムはこうだから食べよう」、「こういう理由で体に良いからとろう」と、決める傾向が強いようです。

疫学とは何でしょう。これは、疫学辞典の冒頭に書いてある文章を少し短くしたものです。

「明確に規定された人間集団に出現する健康関連のいろいろな事象の頻度と分布（記述疫学）、およびそれらに影響を与える要因を明らかにして（分析疫学）、健康関連の諸問題に対する有効な対策樹立に役立てるための科学」と定義されています。

簡単にいってしまえば、「これは人（集団）で起こるか」、「それは現実的に意味があるか」、有

意か否かではなく、現実的に意味があるかを探ることのほうが疫学者の役割であると私たちは考えています。

メカニズムを解くのではなく、量を決めることが疫学の役割です。そして、その量のことを最近ではエビデンスと呼ぶことがあります。しかしながら、エビデンスの語源は量だけとは限りません。私個人としては、エビデンスという言葉の誤用が、現在、社会で問題になりつつあると考えています。

さらに、「なぜ」のメカニズムは機能(function)を表しますが、それに対して、エビデンスといわれる量の科学はeffectを示します。すなわち、functionの科学とeffectの科学の違いといえるでしょう。

「これを食べよう」や「これは避けよう」を医学の上で決める時には、先に「発見の科学」の研究を行います。そしてメカニズムを解明し、メカニズムの解明がほぼできたところで「検証の科学」を行います。これが疫学研究です。

そして、これくらい食べた場合、これくらいたばこを吸った場合、この病気の発症はほぼ無視してもいいという量が決まってくるわけです。これは異なる科学であり、両方とも双方的な役割を持っているとお考えください。片方が片方の代理を務めることはできません。

疫学データは「結果を読むな！ 方法を読め！」が大切

ここからは論文の読み方、特に人の栄養に関する疫学のデータをどう読むかについてお話し

したいと思います。

大切なのは、「結果を読むな！ 方法を読め！」

ということです。これは、それぞれの専門家なら必ずいう言葉で、疫学者も同じことをいいます。疫学研究の結果を読む前に方法をきちんと読んで、その方法が正しくないものは、結果は読まないほうがいいということです。

そしてその次にくるのが、「生物学的にあり得るか？」です。これはメカニズムの裏打ちがあるかを意味します。メカニズムの裏打ちがなかったら、それはたまたま出た結果かもしれないと考え、読まないほうがいいのです。その一方で、まだメカニズムの科学者が気づいていない未知のものが疫学的に現象として見えてしまった可能性もあります。だから、ひょっとしたら金鉱を発見するかもしれない。すなわちゴミか金かはわかりませんということです。

最後に、生物学的にはあり得る「バイオロジカル フロードビリティー」である可能性も否

めません。方法は正しかったという場合は、「価値あり」として世の中にその情報を広めることになります。

ですから、大学では疫学の結果ではなく、疫学の方法論の講義をすることになります。ところが、薬剤以外の疫学の方法論の講義は、日本の大学で最近までほとんどされてきませんでした。従って読める人が非常に少ない。それどころか、できる人が少ない。ということは、怪しい結果も導き出される可能性があることになります。

では、そういう変なものもあるかもしれない玉石混交の結果を、どうしたらきちんと見極められるのか。それを考えたのが系統的総説、Systematic reviewです。その Systematic review を数量的に行うことをメタ・アナリシスと呼びます。

メタ・アナリシスが導く「エビデンスに基づく医療」

古い例を紹介します。これは薬剤疫学の領域ですが、「エビデンスの蓄積」と「専門家の勧告」のズレに関する研究論文で、対象は急性心筋梗塞です。これは心臓の冠動脈に急激な狭窄や閉塞が起こり血流が途絶える病なので、その血栓を溶かせばいいわけです。つまりは血栓溶解剤が緊急の治療法となります。図表5は急性心筋梗塞患者に対する血栓溶解剤は効果があるかどうか、患者さんを使ったトライアルで、この研究を時系列で並べてあります。

最初の研究は1959年に行われました。23人の患者に血栓溶解剤を用い、Odds Ratio (オッズ比) が0.5と書いてあるので血栓溶解剤は有

効であった、死亡率は0.5であったと読みます。ということは、逆数をとると2倍助かったということになります。

問題はその次です。1963年ころに2回目の研究があり、今度は42人の患者さんで行われました。前回の23例に42例を足して65例になったもので、累積のメタ・アナリシスは Cumulative meta-analysis と呼ばれています。そうすると人数が増えるので、研究は徐々に緻密になっていき、信頼区間は狭くなります。

ところが2番目、3番目までは血栓溶解剤は有効らしいが1を上回っているので有意差がない。従って、この時点で血栓溶解剤は有効とい

2

う判断はできませんでした。

そこで4番目の研究が行われます。患者さんの数はかなり増えましたが、血栓溶解剤はワンポイントで見てもほぼ1に近い。すなわち役に立たない。当然ながら95%信頼区間は非常に広く、役に立たないというがっかりの結果でした。

1970年の初期にさらに研究を行い、1000人を上回る患者さんを用いましたがオッズ比が0.8くらいで、ここでも優位ではありませんでした。

2000人を超えた時初めてオッズ比が0.7から0.8になり、95%信頼区間が1を下回りました。統計学的に有意に治療効果があり、その治療効果はおよそ0.7~0.8の間のオッズ比であるということがわかったのです。

それからはどんどん患者さんが投入され、すなわち命が投入されて研究が行われましたが、少しだけ1のほうに寄っていくものの、オッズ比はほとんど動きません。1990年に至るまでに計4万8154人の患者さんが投入され、この時代に至るまで研究が繰り返されました。

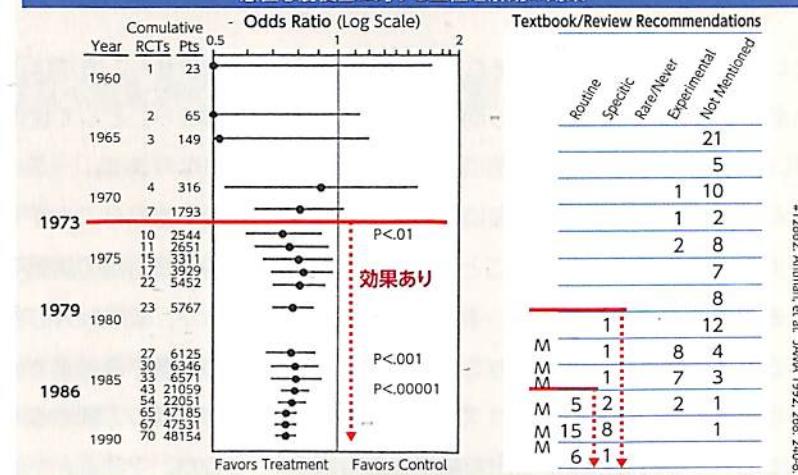
この研究における教訓について考えてみましょう。この研究論文は非常に有名なのですが、中央のグラフよりも右側のほうがすごいことをやってくれました。それは、この研究論文がその時代時代の専門医が使うマニュアルに載っているか、教科書に書いてあるか、ガイドラインに書いてあるかをつぶさに見たことです。

すると、1973年ころに有意になり、1975年ころで有意な研究が何回かあるわけです。にもかかわらず、1980年になるまで現場の医者が使える書物に出てこなかったことが明らかになりました。要するに、研究が現場に伝えられなかつたのです。そしてその後もどんどん患者さんが投入され、命が無駄にされたのではないかという反省が起こりました。

しかしながらもう1つ、私たちはここから貴重なことを得たのです。中央のグラフの時系列を見ると、1959年の最初の研究はオッズ比0.5で、要するに治療効果が大きい。ところがそれが真ん中に寄ってきて、最後は有意だけが0.5より小さい0.78ぐらいで收まります。こういう現象がよく起こります。

もう少し詳しくいうと、最初は非常に効果が強いか非常に悪いという論文が出て、しばらくそういう論文が続きます。すると、3つ目か4つ目にそれを否定する論文がポーンと出て、また世間を騒がせるわけです。この場合は、4つの論文になります。そして、5つ目以降は落ち着いていきます。これが疫学でよく見られる現象です。

図表5「エビデンスの蓄積」と「専門家の勧告」との比較



私たち疫学者は1つ目の論文はバイチャンスと呼び、2つ目の論文で「そろそろ論文を読もうか」という気になり、3つ目の論文が出ると「これは真面目に取り組まないといけないな」という気持ちになる。そのころになると否定の論文が出てきて、「やはりそういうものもあるか」と思うわけです。これがメタ・アナリシスの心です。

従って、メタ・アナリシスはかなりの数の研究論文が出ないとできませんが、先のケースですと、1980年以降の研究は必要なかったことになります。すなわち、メタ・アナリシスを行うことで研究のやり過ぎを予防し、少ない命で多くの人を助けようとする動き、すなわち Evidence-based medicine (EBM) 一根基に基づく医療への流れにつながります。

せっかく集めた論文も疑わしければ除外する

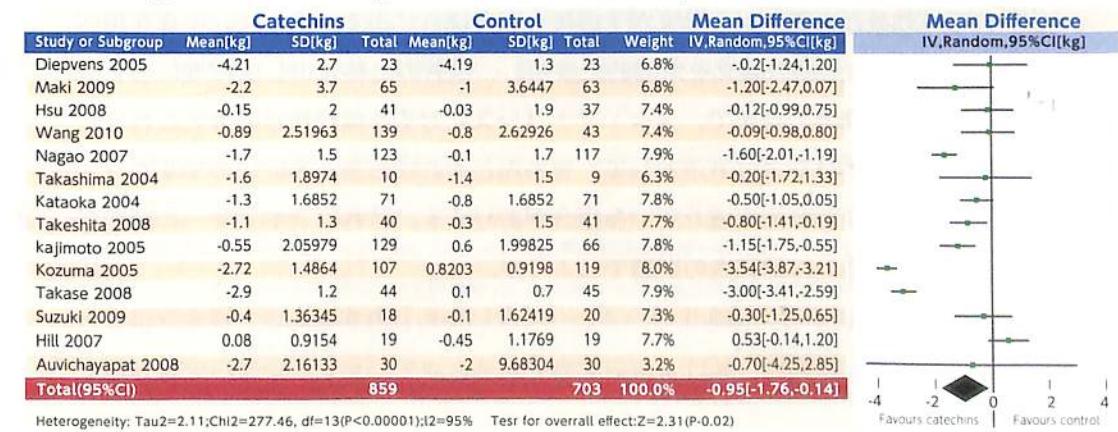
英國に「コクラン共同計画」というメタ・アナリシスの団体があります。「肉と大腸がん」や「EPAとうつ」などのテーマを対象に、これまでに発表された世界中の医学系研究論文を集め一挙に評価しようというNPOで、アーチボルト・コクラン医師の提唱で始まりました。

コクラン共同計画が、緑茶カテキンで体重が減るかどうかについての研究論文を集めました。緑茶の研究でしたら当然日本にあるわけで、相当数が集まつたそうです。図表6の右端のグラフの右側がFavours controlですから、体重増

を示しています。左側はFavours catechinsですから、カテキンが入っている緑茶で体重が減っています。黒いひし形の端から端までが95%信頼区間です。

カテキンが入っている緑茶のほうが有意に体重を下げました。体重1kgですが、有意に減るという結果が出ています。ところが著者は、「緑茶調製品は、過体重あるいは肥満の成人において、統計学的に有意でないわずかな体重減少を誘発するようである。体重減少量が少ないため、臨床的に重要ではない」と結論付けました。体

図表6 Weight loss (sorted highest to lowest total daily dose of catechins)



Heterogeneity: $Tau^2 = 2.11$; $Chi^2 = 277.46$, $df = 13$ ($P < 0.00001$); $I^2 = 95\%$ Test for overall effect: $Z = 2.31$ ($P = 0.02$)

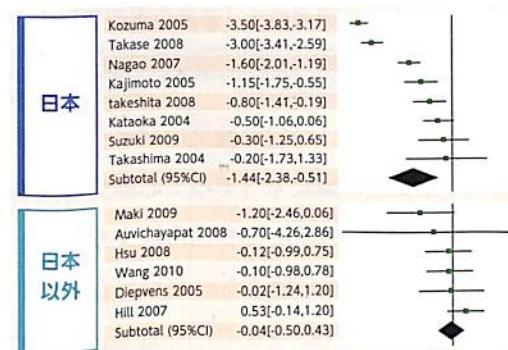
図表7 日本の研究を除外して計算すると

重減少を誘発するのだから重要だと思うのですが、結論は「重要ではない」でした。

なぜでしょう。それは、この論文には付帯条項が付いていたのです。図表6は主条項です。付帯条項は、日本の研究と日本以外の研究の2つに分けました(図表7)。上が日本で、下が日本以外です。この結果のはらつきをごらんになれば、おわかりいただけるでしょう。日本の研究では研究間のはらつきが大きく、日本以外の研究では研究間のはらつきが小さいのです。

この論文には、「日本という一国の中で行われた研究で、これほど結果がばらつくのは何かおかしい。しかし何がおかしいのか私たち読む側にはわからない。従って、日本で行われた研究をすべて除外することにした」と書いてあります。そして、日本を除いた6つの研究で最終結論がつくれました。

1つの論文ではわからないが、多数の論文を



集めることによって結論をあぶり出すメタ・アナリシスには数多くの功があります。しかしながらさまざまな偏りが生じるという罪の一面があることも事実です。

論文の読み方は極めて難しい。研究側にとっても難しいし、読者側にとっても難しいだろうと考えています。功も罪も含め、その論文を読むか読まないか、その論文をどう読むかは読者側の問題です。いろいろなところにメタ・アナリシスや栄養の疫学研究があります。この機会に疫学の研究論文をどう読むべきか、じっくりと考えていただければ幸いです。

討議の抜粋

(敬称略)

大樹 ヨーロッパでは2004年に治験の論文審査が厳格化され、これまで効果があったとされるコレステロール降下剤が、2005年以降はすべて効果がないという結果になりました。ところが各治験では全く有意差がないのに、なぜかメタ・アナリシスでは効果があるという論文がその後ちらほらと出ていました。中身を調べると、2004年以前の論文を意図的に多用し、最新の治験は申し訳程度にちょっと入れていた作為の結果で、メタ・アナリシスに対する信頼性が随分失われました。メタ・アナリシスは方法として、きちんとやればもちろんいいのですが、かなり恣意的な使われ方をされている場合があります。これはスポンサーの問題でしょうか、あるいは研究者の問題でしょうか。

佐々木 ここで言葉にするのはあまりふさわしくないかもしれません、資金提供者と、もう1つ大きな問題は、私は企業研究者であると見ています。企業研究者が論文

作成グループに加わっている場合、当然その企業の開発に利する研究をいたしますので、バイアスがかかって仕方がないところでしょう。従って、クライテリアやそのプロセスを明文化すると同時に、所属と、その人のコントリビューションをきちんと論文に明記することが、最近厳しく問われるようになりましたが、全くそのとおりだと思います。

清水 企業の研究者が入ることによって、データにバイアスがかかるることは確かにあります、例えば緑茶カテキンについて、日本人特有の食生活、腸内細菌叢、遺伝的な特性、風土など日本独自の特殊性もあると思うのですが、それらが影響してくる可能性を排除できるのでしょうか。

佐々木 それは排除できません。では、なぜコクランレビューが緑茶カテキンの研究で日本の論文を除外したかといいますと、それは日本人で有効かを見たかったのではなくて、純粋、単純、シンプルなメタ・アナリシスとして有効か否かを見たかったからで、民族性を超えた結果を出したいというのが彼らの目的であったと読めるのではないかと思います。

柴田 先生がお話しされた医学方法論、疫学について、Evidence-based medicineという言葉だけは先行していますが、実際に講義をしている大学は非常に少ないですね。日本の医学教育の中で、公衆衛生学教室は、もっと指導的な立場に立たなければいけないと思いますが、いかがでしょうか。

佐々木 公衆衛生学は、医学部の中では本当に弱小なのです。東京大学は、11年前に公衆衛生大学院(School of Public Health)を立ち上げ、私はその立ち上げとともに赴任しましたが、すでに存在していたのは京都大学だけで、東京大学は2番目でした。ところが、その当時アメリカには40ありました。日本ではどうしても薬剤疫学がマジョリティになりますが、欧米では本来の予防医学や感染症コントロールの疫学が中心になるべきとの流れでした。私もその流れに沿って、やはり日本中で、きちんとした量的・質的な疫学ができるようにと願っています。

■ささき・さとし 昭和56年、京都大学工学部資源工学科卒業。その後、京都大学大学院工学研究科修士課程中退。平成元年、大阪大学医学部医学科卒業。平成6年、大阪大学大学院医学研究科博士課程修了。医学博士(公衆衛生学)。平成6年、ベルギーのルーベン大学大学院医学研究科博士課程修了。名古屋市立大学医学部公衆衛生学教室助手を経て、平成8年、国立がんセンター研究所支所臨床疫学研究部室長。平成14年、独立行政法人国立健康・栄養研究所栄養所要量策定企画・運営担当リーダー。平成18年、同研究所栄養疫学プログラム・プログラムリーダー。平成19年から現職。「佐々木敏のデータ栄養学のすすめ」(女子栄養大学出版部2018年刊)、「佐々木敏の栄養データはこう読む!」(女子栄養大学出版部2015年刊)など著書多数。