

## 4 暴露疫学等調査

### 総論 - 疫学研究の方法 -

疫学研究は、人間集団に起こっている事象を確率・統計を手技に用いて客観的に示す方法論である。この方法論を“化学物質の人への健康影響（疾病の発症、あるいは、機能の低下）”という問題解決のために適用する事により、“化学物質の暴露量が多い人達が、少ない人達に比べて、ある病気になる確率が高いか否か”についての検証を行う事になる。以下では、疫学研究で用いられる、疾病頻度の指標、暴露の指標、暴露要因と疾病の関連性の指標、及び研究デザインの概略について解説した上で、暴露要因と疾病の因果関係を評価する際の方法論の現状を概括する。

#### 1. 疾病頻度の指標

暴露要因と疾病との関連を明らかにするためには、人口集団から発生する疾病の頻度を測定することが基本になる。最も一般的な疾病頻度の指標は、罹患率（incidence）、死亡率（mortality）、及び有病率（prevalence）である。

#### 罹患率と死亡率

罹患率と死亡率は、以下のように表される。

$$\text{罹患率} = \frac{\text{集団から一定期間に新たに発生する患者の数}}{\text{集団の観察人年}}$$

$$\text{死亡率} = \frac{\text{集団から一定期間に発生する死亡者の数}}{\text{集団の観察人年}}$$

観察人年（person-years）は、1人を1年間観察すれば1人年に相当する。2人を5年間観察すれば10人年、5人を2年間観察しても10人年に相当する。がんなどのまれな疾病の罹患率や死亡率は、通常、10万人年あたりの新規発生患者数や死亡者数で表現される。例えば、ある1万人の集団を10年間観察したところ160人の胃がん患者が発生したとすれば、この集団における胃がんの罹患率は160（対10万人年）となる。

罹患率や死亡率の分母として、観察人年ではなく、観察開始時の集団の人数を用いる場合がある。これを累積罹患率・累積死亡率と呼ぶ。例えば、1万人の集団を10年間観察したところ160人の胃がん患者が発生したとすれば、この集団における胃がんの10年累積罹患率は160人/1万人＝1.6%となる。

## 有病率

有病率は、以下のように表される。

$$\text{有病率} = \frac{\text{ある一時点ですでに疾病にかかっている患者の人数}}{\text{集団の人数}}$$

罹患率が集団の観察期間を考慮に入れた指標であるのに対して、有病率はある一時点での有病者の割合を示している。10万人を対象に胃の内視鏡検査を実施したところ、10人に胃がんを認められたとすれば、この集団における胃がんの有病率は10人/10万人となる。

暴露要因と疾病との関連を研究するためには、罹患率を指標として用いることが、通常は最も適切である。罹患率が使えない場合には、次善の手段として、死亡率や有病率が用いられる。

## 2. 暴露の指標

喫煙の健康影響を疫学研究において検証する際に、暴露の指標は比較的明瞭である。喫煙者・非喫煙者・過去喫煙者、あるいは、一日当たりの喫煙本数と喫煙年数を組み合わせて、喫煙量を定量的に測る事も出来る。一方、化学物質暴露の指標については、シエチルスチルベストロール (DES) の服用など一部を除いて、目に見えるものではなく、かつ、極めて微量、また、様々な暴露源が存在するために簡単ではない。

最もわかりやすい指標は、血液・組織・尿などの生体試料中の存在量であり、吸収・代謝・排泄において個人差がなく、恒常性の影響を受けなければ、どの程度の期間の暴露量を反映しているかにもよるが、様々な暴露源からの暴露をあらゆる最も良い指標となる。しかしながら、化学物質の暴露によって疾病が生じるという時間関係の整合性を保つためには、疾病発症以前の生体試料を用いなければならないため、後述するコホート内症例対照研究において保存されている血液などを用いて化学物質を定量する事が、現実的には最も理想的である。ただ、生体試料中化学物質濃度がその疾病の存在による影響を受けないという仮定が成立する場合には、症例対照研究や断面研究によっても、化学物質と疾病との関連をある程度評価可能である。例えば、胃がん罹患患者であれば、その症状により食習慣が変化した結果としての血中の化学物質濃度を反映している可能性が高いが、乳がん罹患患者であれば、そのような変化は少ないものと仮定出来、乳がん罹患以前の血中化学物質濃度の良い指標であると見なせ得る。

また、化学物質の暴露量と健康影響は必ずしも直線関係ではなく、閾値の存在や量による健康影響の方向性のシフト（低濃度では良い健康影響をもたらすが、一定濃度以上で悪影響を示す）などの可能性が考えられるので、定性的ではなく、定量的評価が望まれ、その意味でも生体試料中の存在量による暴露の評価は、人における健康影響を知るうえにおいて重要である。

化学物質暴露を生体試料などで直接測定出来ない場合には、化学物質暴露をあらゆると思われる代理指標が用いられる。化学物質を扱っている職業、塩素系農薬の使用状況、化学物質を含む食品の摂取量、化学物質が溶出されると思われる食器や玩具の使用状況などが挙げられるが、いずれも間接的に化学物質暴露を示す指標に過ぎないという限界はある。

### 3. 関連性の指標

個人が一定期間に疾病に罹患する確率や、疾病で死亡する確率を、リスク (risk) という。暴露要因と疾病との関連性を調べる際には、暴露要因がある場合とない場合（または暴露量の多い場合と少ない場合）で疾病リスクを比較する。暴露要因と疾病の関連性を示す代表的な指標には、相対危険度(relative risk, RR)、オッズ比(odds ratio, OR)がある。相関係数(correlation coefficient)も関連性の指標として使われる場合がある。

#### 相対危険度

相対危険度は、暴露要因がある場合とない場合の疾病リスクの比である。本来は二つの集団における累積罹患率の比であるが、一般的には、二つの集団における観察人年当たりの罹患率の比であらわされる。例えば、喫煙の肺がん罹患率が400（対10万人年）で、非使用者の罹患率が100であれば、喫煙による肺がん罹患に対する相対危険度は $400/100=4$ となる。

※ 喫煙者 10,000 人と非喫煙者 10,000 人を 10 年間追跡（コホート研究）

		肺がん発症	罹患率
喫煙者	10,000	400	400/10万人年
非喫煙者	10,000	100	100/10万人年

また、有機塩素系農薬の暴露量が多い集団の乳がん罹患率が300で、少ない集団の罹患率が100であれば、相対危険度は $300/100=3$ となる。相対危険度は、二つ以上の集団の罹患率を直接測定することができる無作為割付臨床試験や前向きコホート研究で使われる。

#### オッズ比

オッズ比も、暴露要因がある場合とない場合の疾病リスクの比である。オッズ比は、罹患率を直接測定できない症例対照研究や、罹患率ではなく有病率を疾病頻度の指標とする断面研究で用いられる。無作為割付臨床試験や前向きコホート研究における相対危険度に相当する。例えば、肺がん症例100人と対照100人中に、喫煙していた者がそれぞれ80人と50人いた場合を想定する。

※ 肺がん患者 100 人と対照 100 人の過去の喫煙歴

	肺がん患者	対照
	100	100
喫煙者	80	50
非喫煙者	20	50

肺がん患者における喫煙ありのオッズは $80/20=4$ 、対照におけるそれは $50/50=1$ となり、両オッズの比（オッズ比）は、 $4/1=4$ となる。相対危険度は、喫煙者からの肺がん罹患率と非喫煙者からの肺がん罹患率の比であるべきであるが、単純に上記の表を横方向に見ることは出来ない。実際は、100人の肺がん患者が発症する一方では、100人の肺がんでない対照の何十倍もの肺がんを発症していない人達が存在する。したがって、喫煙者からの肺がん罹患率は、

$80/(80+50k)$  ( $k$  は定数)、非喫煙者からの肺がん罹患率は、 $20/(20+50k)$  と仮定する事が出来、相対危険度は、 $\{80/(80+50k)\}/\{20/(20+50k)\}$  となる。 $k$  が十分に大きい数だとすると (実際に、数十倍から数百倍)  $80+50k$  や  $20+50k$  は共に  $50k$  としても無視出来る差しか存在しないので、分子と分母の  $k$  が相殺され、 $(80/50)/(20/50)$  となりオッズ比  $(80/20)/(50/50)$  と同じ式となる。すなわち、稀にしか発症しない疾病に関しては、症例対照研究におけるオッズ比は、コホート研究における相対危険度の良い推定値を与え得る。

### 相関係数

相関係数は、断面研究や地域相関研究で用いられることのある関連性の指標である。相関係数は  $-1$  から  $+1$  の範囲の値をとり、 $+1$  ( $-1$ ) であれば二つの変数に完全な正の(負の)直線的関係があることを示し、 $0$  であれば二つの変数に何ら直線的関係がないことを示す。

## 4. 研究デザイン

特定の暴露要因と疾病との関連を疫学的に調査する場合、研究方法はただ一つではなく、複数の方法が存在する。これらの研究方法を「研究デザイン」(study design)と呼ぶ(表2-1)。代表的な研究デザインには、次のようなものがある。

- 1) 無作為化割付臨床試験 (randomized controlled trial)
- 2) 前向きコホート研究(cohort study)
- 3) 後ろ向きコホート研究 (retrospective cohort study)
- 4) 症例対照研究(case-control study)
- 5) コホート内症例対照研究(nested case-control study)
- 6) 断面研究(cross-sectional study)
- 7) 地域相関研究(ecological study)

無作為化割付臨床試験では、研究者が積極的に治療法や予防法を行う (介入する) ため、「介入研究」(intervention study) と呼ばれることがある。それ以外の研究デザインでは、研究者は積極的な介入を行わず、対象者自身の日常的な食生活などの行動を調査する (観察する) ことで研究を行うため、「観察研究」(observational study) と呼ばれることがある。

### 1) 無作為割付臨床試験

無作為化割付臨床試験では、まず、乱数表などの手段を使って対象者を「介入群」と「対照群」に振り分ける。この措置を「無作為割付」(random allocation) という。介入群に割り付けられた対象者には、疾病予防効果を評価しようとする栄養素等の化合物を補給剤 (supplement) として投与する。対照群に割り付けられた者には、補給剤と見かけは同じだが化合物の含まれないプラセボ (placebo) を投与する。その後、介入群と対照群に対して追跡調査を行い、疾病の罹患率を比較する。介入群の罹患率が対照群よりも低くなれば、化合物の疾病予防効果が示されたことになる。無作為化割付臨床試験では、研究者の恣意や対象者の希望を考慮せずランダムに介入群と対照群に振り分けるので、二群の特性をそろえることができ、介

入の効果をより正しく評価することができる。そのため、無作為割付臨床試験は、他の研究デザインと比べて、最も妥当性が高い研究方法と考えられている。

内分泌系に作用することが試験管内実験などで示されている化学物質については、ジエチルstilbestrol（DES）の流産予防の有効性評価のために過去に実施された無作為割付臨床試験において、追跡期間を延長してがんの罹患について把握した研究や、豆乳の健康影響を評価するために実施された無作為割付臨床試験など、介入により良い健康影響を期待して行われた研究が該当するが、悪い健康影響を評価する研究は倫理的に存在し得ない。

## 2) 前向きコホート研究

前向きコホート研究では、健康人の集団を対象として、暴露状況を調査票などの手段で調査する。この調査対象集団を「コホート」(cohort)という。この集団を長期間追跡して、疾病の罹患率を調査する。その上で、はじめに調査した暴露状況と、その後の罹患率の関係を分析する。

前向きコホート研究では、対象者が疾病に罹患する以前の暴露状況を調査するので、症例対照研究で問題となるような思い出しバイアスを回避できる。前向きコホート研究ではまた、交絡要因に対する情報を最初に収集しておけば、データ解析の段階で統計的な補正を行うこともある程度可能である。一方、この研究方法では、多数の集団（数万から数十万）を長期間（数年から数十年）追跡しなければならないため、多大な手間と費用がかかるのが欠点である。また、情報が収集されていない交絡要因や未知の交絡要因については、その影響を統計的に補正することはできない。

## 3) 後ろ向きコホート研究

後ろ向きコホート研究では、すでに暴露が起ってしまった後で、研究者が事後的に（後ろ向きに）その状況を調べ、さらにその集団を追跡調査することで、疾病の発生を確認する。事故によって高濃度の化学物質に暴された産業施設労働者の暴露状況を、事故が生じた後で調査し、その集団からのがん発生や死亡を追跡調査によって明らかにする場合などに、この研究方法が用いられる。

後ろ向きコホート研究では、疾病頻度に関する指標として、**O/E比** (observed to expected ratio)、**標準化死亡比** (standardized mortality ratio, SMR)、**標準化罹患比** (standardized incidence ratio, SIR) が使われる。

O/E比は、以下のように表される。

$$\text{O/E比} = \frac{\text{調査集団で実測された (observed) 疾病の頻度}}{\text{調査集団と性別や年齢構成が同じ一般の人口集団から期待される (expected) 疾病の頻度}}$$

O/E比が1より大きければ、調査集団から生じた疾病の頻度（罹患数や死亡数）が、一般集団から期待される頻度より大きいことになるので、暴露要因による疾病リスクの上昇を示すことになる。

標準化死亡比と標準化罹患比は、O/E比と類似の指標であり、死亡数や罹患数の実測値と期待値が等しいときには100で表す。これらの指標が100より大きければ、暴露要因による死亡率や罹患率の上昇を示すことになる。

後ろ向きコホート研究では、対象者に暴露がおこってから、その後に生じた疾病との関係を調査するので、暴露要因と疾病の時間的前後関係を正しく評価できる。その一方、すでに暴露がおこった後で研究を始めるため、暴露の程度を定量的に評価することが困難な場合があり、交絡要因に関する情報も十分に収集できない場合が多い。したがって、例えば、特定の化学物質に暴露した産業施設労働者を対象にこの方法で調査を行い、がん死亡の実測値が一般の集団からの期待値より高いことが観察されたとしても、それが化学物質そのものの影響なのか、それ以外の物質の影響なのかを、区別することは通常難しい。

#### 4) 症例対照研究

症例対照研究では、最初に、すでに疾病にかかった人を「症例」(case)として選び出す。次に、この症例と性別や年齢などの要因が似た人を「対照」(control)として選ぶ。対照は、症例と同じ地域に住む健康な住民から選ばれる場合と、症例と同じ病院に入院している患者から選ばれる場合がある。前者を「住民対照」(population control)といい、後者と「病院対照」(hospital control)という。対照を選択する際には、性別や年齢などの特性を、対応する症例とそろえることが多い。これを「マッチング」(matching)という。マッチングは、症例と対照の個々の組合せごとに特性をそろえる場合と、症例全体と対照全体で特性の分布をそろえる場合がある。前者を「個人マッチング」(individual matching)、後者を「頻度マッチング」(frequency matching)という。

症例対照研究では、調査票などを用いて、過去の暴露状況を思い出してもらい、症例と対照で比較する。例えば、ある化学物質の過去の暴露量が、対照より症例で多ければ、そのEDCは発がんを促進する作用を持つ可能性がある。

症例対照研究は、無作為割付臨床試験や前向きコホート研究より少ない対象者(数百名程度)で研究を行えるという長所があり、追跡調査を行う必要もない。データを分析する際に交絡要因の影響を統計的に補正することも、ある程度可能である。

その反面、過去の暴露状況を思い出してもらって調査するため、記憶に偏りが生じ、暴露と疾病の関連を過大評価したり、反対に過小評価したりする危険性がある(思い出しバイアス)。また、住民対照を選択する際に、一般住民の中でも特に調査に協力的で生活習慣も健康的な者を偏って選んだり、病院対照を選択する際に、一般の住民より生活習慣の不健康な者を選んだりして、暴露と疾病との関連を正しく評価できない危険性がある(選択バイアス)。一般に、症例対照研究の結果の信頼性は、無作為割付臨床試験や前向きコホート研究などと比べ、相対的に低いと考えられている。したがって、症例対照研究で暴露要因と疾病の関係を評価する時には、結果の解釈に十分な留保が必要になる。

#### 5) コホート内症例対照研究

コホート内症例研究は、前向きコホート研究の参加者を用いて行なわれる。参加者の中から、追跡調査期間中に特定の疾病にかかった者の全員を症例として選び、それ以外の参加者の一部から対照を選び、症例対照研究としての分析を行う。コホート集団を設定した時に採取し保存

しておいた生体試料（血液検体など）を測定して、症例と対照を比較する場合にこの方法が使われる。例えば、コホート集団から発生した乳がん症例を選び出し、さらに性別や年齢をマッチさせた対照を選んで、凍結保存しておいた血清中のPCB濃度を測定する。PCBの血清濃度が対照より乳がん症例で高ければ、PCBの乳がんリスク上昇作用が示唆されることになる。

前向きコホート研究の参加者全員について生体試料を測定すると費用がかかるが、症例や対照として選んだ一部の参加者の生体試料だけを分析すれば、安価に研究を行える。また、この方法では、対象者が疾病に罹患する以前に生体試料を採取し保存しておくので、暴露要因（生体指標）と疾病の時間的前後関係を正しく評価することができる。また、交絡要因に対する情報を最初に調査票などで収集しておけば、データ解析の段階で統計的な補正を行うこともある程度可能である。

一方、多数の集団（数千から数十万）を長期間（数年から数十年）追跡しなければならないという前向きコホート研究に共通の問題点に加えて、多数のコホート参加者から生体試料を採取し、後の分析に備えて凍結保存しなければならないため、多大な手間と費用がかかる。また、情報が収集されていない交絡要因については、その影響を統計的に補正することはできない。

## 6) 断面研究

断面研究では、暴露要因と疾病の有無（有病率）を同時に調査し、両者の関連を調べる。例えば、ダイオキシンの血清濃度と子宮内膜症の有病率を同時に調査したところ、ダイオキシンの血清濃度が高い集団の方が、少ない集団より、子宮内膜症の有病率が高ければ、ダイオキシンによる子宮内膜症リスクの上昇が示唆されることになる。

断面研究は、暴露要因と疾病とを同じ時点で（断面で）調査するので、無作為割付臨床試験やコホート研究のような追跡調査を行う必要がなく、比較的簡単に研究を行える利点がある。

反面、この方法では、暴露要因（原因）と疾病の有病率（結果）を同時に調査するので、両者の関連が認められたとしても、その時間的前後関係を正しく評価できない場合がある。一般に、断面研究の結果の信頼性は、無作為割付臨床試験や前向きコホート研究などと比べ相対的に低いと考えられている。したがって、断面研究で暴露要因と疾病の関係を評価する時には、結果の解釈に十分な留保が必要になる。

## 7) 地域相関研究（エコロジカル研究）

地域相関研究では、複数の人口集団（国・都道府県・地域など）を対象として、暴露要因の平均値と、疾病の罹患率や死亡率との関連を調査する。それぞれの集団におけるこれらのデータは、特定の研究のために新しく調査される場合もあるが、別の目的で収集された既存の資料を用いることも多い。地域相関研究は、個人ではなく集団を分析の単位としており、集団全体における平均的な暴露状況や、集団の罹患率・死亡率が指標として用いられるのが特色である。

地域相関研究には、行政的な目的で収集された既存の資料を用いる場合には、比較的簡便に研究を行える長所がある。一方、集団を対象とした研究の結果をそのまま個人に適用できるとは限らないという問題点がある。また、例えば、ある要因と疾病死亡率との正の相関が認められたとしても、それがその要因そのものの影響なのか、別の要因の影響を見かけ上反映しているだけなのかを区別することが難しい。地域相関研究では、こうした交絡要因の影響を制御することが困難な場合が多い。一般的に、地域相関研究の結果の信頼性は、無作為割付臨床試験

や前向きコホート研究などと比べ、相対的に低いと考えられている。

## 5. 研究デザインにもとづく疫学研究の評価

これまで記してきたように、暴露要因と疾病に関する同じ仮説を評価する場合でも、研究デザインによって、調査の実施の容易さや、研究結果の信頼性が異なる。一般に、実施が困難な研究ほど結果の信頼性が高く、反対に、実施が容易な研究ほど結果の信頼性は低い傾向がある。すなわち、研究の実施は、無作為割付臨床試験が最も困難で、前向きコホート研究やコホート内症例対照研究がそれに次ぐ。地域相関研究や断面研究の実施は相対的に容易である。後ろ向きコホート研究や症例対照研究は、これらの中間に位置する。一方、結果の信頼性は、無作為割付臨床試験が最も高く、前向きコホート研究やコホート内症例対照研究がそれに次ぐ。断面研究や地域相関研究の信頼性は相対的に低く、後ろ向きコホート研究や症例対照研究はこれらの中間に位置する。そのため、同じ仮説についての研究が複数存在し、それぞれの結果が一致しない場合、無作為割付臨床試験、前向きコホート研究、コホート内症例対照研究の結果は、他の研究デザインによる結果よりも重視して解釈される。

## 6. 因果関係の評価

### 1) 研究結果の競合的解釈

疫学研究の重要な目的の一つは、暴露要因と疾病の因果関係を明らかにすることである。しかし、例えば、ある化学物質の暴露量が多い集団の疾病の罹患率が、暴露量が少ない集団の罹患率より高いことが特定の研究で観察されたとしても、その結果から即座にその化学物質と疾病との因果関係が明らかにされたと考えることはできない。

疫学研究には、偶然 (chance)、バイアス (bias)、交絡 (confounding)、因果性 (causality) という四つの要因が影響をおよぼす。この四つの要因を研究結果の競合的解釈 (alternative explanations) という。

偶然とは、測定値のランダムな変動によって、暴露要因と疾病の関連についての観察が影響を受けることをいう。バイアスとは、暴露要因と疾病との実際の関連を過大評価したり過小評価したりして、誤った研究結果を導いてしまう現象をいう。交絡とは、暴露要因と疾病の実際の関連性が、第三の要因の影響によって過大評価ないし過小評価されてしまう現象をいう。

特定の研究の結果を解釈する際には、因果性以外の三つの要因がどの程度影響を与えているかを吟味する必要がある。また、疫学研究を行う際には、因果性以外の要因の存在をあらかじめ見込んだ上で、その影響をできるだけ制御できるように調査を設計し、データを分析することが必要になる。

ただし、ここで重要なのは、たとえどれほど優れた研究であっても、三つの要因が全く入り込む余地のない「完全」な研究は存在し得ないという点である。疫学を始めとする実証科学では、理論の正しさ (因果性) を議論の余地なく「証明」することは不可能であり、非実証科学である数学とは、この点で異なっている。

それでは、不完全な研究の知見に基づいて、暴露要因と疾病の因果関係を評価するためには



どうすればよいか。これまで、この「因果推論」(causal inference)の問題に対するアプローチとして、英国の疫学者 Hill による判定規準、ダイオキシンの健康影響に関する全米科学アカデミーの判定規準、化学物質等の発がん性評価に関する米国保健省と国際がん研究機関の判定規準、喫煙の健康影響に関する米国保健省の判定規準などが提唱されている。以下では、これらの概要について解説する。

## 2) Hill の判定規準

慢性疾患を対象とする現代の疫学で、因果推論のあり方が最も鋭く問題化した最初の歴史的事例は、喫煙と肺がんの関連をめぐる問題である。1964年に米国保健省は、『喫煙と健康に関する米国公衆衛生局長官報告』と題する報告書を発刊した(U.S. Surgeon General's Advisory Committee on Smoking and Health, 1964)。報告書では、肺がんを中心とする疾病と喫煙に関する観察的疫学研究を総括した上で、喫煙と肺がんの因果関係を肯定した。

この際に用いられた因果関係の判定規準をさらに発展させて定式化させたのが、英国ロンドン大学のヒルの論文『環境と疾病—関連性か因果性か—』である(Hill, 1965)。この論文は、1965年1月14日に英国王立医学会の産業医学部会で行われた講演の記録である。ヒルは、ある暴露要因と疾病の間で統計的な関連性(association)の存在が観察された場合、「どのような条件のもとで、この関連性から因果関係(causation)の存在を判断することが可能か? どのような根拠に基づいてこうした推論を行うべきか?」という問いを立てた上で、「関連性の観察から因果性の存在を結論づける前に吟味すべき9つの観点(viewpoints)」を提示した。ヒルが言及した具体的事例を適宜示して、表2-2に整理した。

ヒル自身は、因果関係を判断する際の必要条件のリストを示そうとしたわけでは必ずしもなかった。むしろその反対に、「因果関係の存在を認める前に従うべき確実で手早い規則」を定めることに対して警告を発していた。実際、これらがあくまでも「観点」にすぎないことを、例外や反対事例にも言及しながら強調していた。しかし、その後出版された疫学の多くの教科書では、ヒルの「観点」のリストが因果関係を推論するさいの「判定規準(criteria)」として提示され、しかもこの規準の全てが因果性を判断する際の必要条件であるかのような形で示された。このような事情を通して、ヒルの判定規準が、疫学における因果推論の典型例と見なされるようになった。

## 3) ダイオキシンの健康影響に関する全米科学アカデミーの判定

米軍は、ベトナム戦争の際に、「エージェント・オレンジ」を始めとする除草剤を使用した。戦後、ベトナムから復員した在郷軍人が、さまざまな健康障害を訴え始めた。その原因として、除草剤(特に汚染物質として含まれていたダイオキシン)への暴露が疑われるようになった。こうした懸念に應えるため、米国議会は、「エージェント・オレンジ法」を1991年に制定した。この法律は、ベトナム戦争で使用された除草剤と、その中に含まれていた化学物質の健康影響に関する報告書を、2年間に1回、10年間にわたって作成し、議会に報告することを求めている。1994年に最初の報告書が公表され、これまで1996年、1998年、2000年、2002年の四回、その後の新しい科学的知見を加えて、報告書が更新されている。報告書の作成は、全米科学アカデミーの医学部会が、委員会を設定して行っている。

2002年の改訂版報告書に示された判定を表2-3に示す。報告書作成委員会は、除草剤と個

別の健康障害との関連性を検討するにあたり、疫学研究のレビューと解釈に労力の大半を費やした。その上で、関連性を示す「十分な知見がある」「限定的／示唆的な知見がある」「不十分／不適当な知見がある」「関連性を否定する限定的／示唆的な知見がある」という四段階の判定を行った。

ここで、例えば、「十分な知見がある」という判定は、除草剤と疾患との関連性が「偶然、バイアス、交絡の影響を相応の信頼性をもって否定できるような研究において」観察されている場合に適用されている。「限定的／示唆的な知見がある」という判定は、両者の関連性についての知見が「偶然、バイアス、交絡の影響を確実に否定できないため限定的である」場合に適用されている。また、「知見が不適当／不十分」という判定は、既存の知見の「統計的検出力が不適当（で偶然の影響を排除できない）」「交絡の影響を制御していない」場合に適用されている。

#### 4) 化学物質等の発がん性評価に関する米国保健省の判定規準

米国保健省 (Department of Health and Human Services) は、1980 年以来9版にわたり、「発がん物質に関する報告 (The Report on Carcinogens)」を出版している。2000 年に公表された最新の第9版には、218 種の化学物質等が、「ヒトに対する発がん物質であることが知られている (Known to be human carcinogen)」または「ヒトに対する発がん物質であると合理的に想定される (Reasonably anticipated to be human carcinogen)」として掲載されている。第9版で用いられている判定規準を、表2-4に示す。「ヒトに対する発がん物質であることが知られている」という判定は、「人での研究により、発がん性に関する十分な知見 (sufficient evidence) がある。それらの研究では、要因、物質、混合物への暴露と、がんとの因果関係が示されている」ような場合に適用されている。一方、これより下位の「ヒトに対する発がん物質であると合理的に想定される」という判定は、「人での研究により、発がん性に関する限定的な知見 (limited evidence) がある。それらの研究では、因果関係が存在するという解釈にも一定の信頼性があるが、偶然、バイアス、交絡要因という競合的解釈の可能性が、適切に排除されていない」ような場合等に適用されている。

#### 5) 化学物質等の発がん性評価に関する国際がん研究機関の判定規準

国際がん研究機関 (International Agency for Cancer Research) は、1972 年以来、種々の化学物質等の発がん性を評価した、「ヒトに対する発がんリスク評価に関する IARC モノグラフ」(IARC Monographs on the Evaluation of Carcinogenic Risks to Humans) を、シリーズで刊行している。これらのモノグラフには共通の方法論が採用されている。すなわち、(a)人と実験動物の発がん性に関する知見の程度の評価、及び、(b)発がん性とそのメカニズムの評価に関連する他のデータの評価に基づき、(c)総合評価を行う。

(a)において疫学研究を評価する際には、表2-5に示すような4段階の判定を用いている。このうち、最上位の「発がん性に関する十分な知見」という判定は、「偶然、バイアス、交絡が、相応の信頼性をもって排除されている研究において、暴露とがんの正の関連性が観察」されており、「物質、混合物、暴露状況とがんとの因果関係が確立している」と判断されるような場合に適用されている。第2の「発がん性に関する限定的な知見」という判定は、「物質、混合物、暴露状況への暴露とがんの正の関連性が観察されており、因果性が存在するという解釈にも一定の信頼性がある」と判断されるものの、「偶然、バイアス、交絡が、相応の信頼性をもつ

て排除されていない」ような場合に適用されている。第3の「発がん性に関する不適切な知見」は、「暴露とがんの因果関係の有無に関する結論を出す上で、既存の研究の質、一致性、統計的検出力が不十分である」ような場合に適用されている。第4の「発がん性の欠如を示唆する知見」は、「人類が遭遇することが知られている全範囲の暴露レベルをカバーする複数の研究が存在し、物質、混合物、暴露状況に対する暴露と、研究対象のがんととの正の関連性が、観察したどの暴露レベルにおいても認められないという点で相互に一致している」ような場合に適用されている。一方、(a)において動物実験を評価する際には、表2-6に示すような基準を用いて、同じ4段階の判定を適用している。

これらを踏まえた(c)の総合評価では、疫学研究と動物実験における知見の程度を組み合わせ、表2-7に示すような4段階の判定を行っている。このうち、グループ1の「ヒトに対する発がん性がある」という判定は、「ヒトにおける発がん性に関して、十分な知見 (sufficient evidence) が存在する」ような場合に適用されている。グループ2Aの「おそらく (probably) ヒトに対する発がん性がある」という判定は、「ヒトにおける発がん性に関して、限定的な知見 (limited evidence) が存在し、しかも、実験動物における発がん性に関して、十分な知見 (sufficient evidence) が存在する」ような場合に適用されている。グループ2Bの「ヒトに対する発がん性の可能性がある (possibly)」という判定は、「ヒトにおける発がん性に関して、限定的な知見 (limited evidence) が存在し、しかも、実験動物における発がん性に関して、不十分な知見 (less than sufficient evidence) が存在する」ような場合に適用されている。グループ3の「分類不能」という判定は、「ヒトにおける発がん性の知見が不適切 (inadequate) で、実験動物における発がん性の知見が不適切 (inadequate) または限定的 (limited) である」ような場合に適用されている。グループ4の「おそらく (probably) ヒトに対する発がん性がない」という判定は、疫学研究と動物実験の双方で「発がん性の欠如を示唆する知見 (evidence suggesting lack of carcinogenicity) が存在する」ような場合に適用されている。

## 6) 喫煙の健康影響評価に関する米国保健省の判定規準

喫煙の健康影響に関する米国公衆衛生局長官の報告書は、第1回目の1964年版以来、ほぼ毎年のように新しい報告が公表されてきた。これらの報告書は、単純な改訂版ではなく、(循環器疾患や慢性呼吸器疾患など) 特定の疾患にフォーカスしたものの、受動喫煙の健康影響にフォーカスしたものの、女性や人種のマイノリティへの健康影響に焦点をあてたものなど、さまざまな内容である。

2004年5月には、通算で28冊目にあたる最新の報告書「喫煙の健康影響に関する米国公衆衛生局長官報告」"The Health Consequence of Smoking: A Report of the Surgeon General" が公表された。この報告書では、1964年版の第1回報告書と同様に、能動喫煙の健康影響について、多数の疾患を取り上げ、包括的な検討を行っている。

報告書は941ページ、全8章からなる。第1章「序論と因果推論へのアプローチ」(Introduction and approach to causal inference) では、因果推論に関する方法論的な考察を行っている。その際、1964年の報告書が提唱した、5項目の判定規準を用いた因果関係の評価法について、「科学的根拠の包括的評価に関して、その後続くモデルとなった」と、その意義を強調している。

その一方で、喫煙と健康に関する米国保健省の一連の報告書では、因果関係の有無に関する結論の部分で用いられた表記法に、ばらつきがあったことを指摘している。例えば、喫煙と動

脈硬化症に関する結論として、報告書によって次のような多様な表現が用いられていた。

「有意な上昇と関連する」(associated with a significant increase) (1969年版報告書)

「危険因子と思われる」(a likely risk factor) (1971)

「主要な危険因子」(a major risk factor)(1973)

「強固な関連性」(strong associations)(1974)

「主要な独立の危険因子」(a major, independent risk factor) (1980)

「最も強力な危険因子」(the most powerful risk factor) (1983)

「原因であり最も強力な危険因子」(a cause and the most powerful risk factor) (1989)

こうした表現の不一致を指摘した上で、因果関係に関する「結論を報告するためのより構造的な枠組み」として、表2-8のような段階的評価を導入した。この段階的評価を行うにあたり、基礎研究と疫学研究の双方が考慮されている。このような段階的評価が取り入れられたのは、喫煙と健康に関する米国公衆衛生局長官の一連の報告書としては、今回がはじめてである。

報告書では、AからDまでの判定をする際に、偶然・バイアス・交絡の影響をどう考慮するかについて、具体的な記述はない。ただし、3)で述べたダイオキシンに関する全米科学アカデミーの報告書と、5)で述べた化学物質の発がん性に関する国際がん研究基金の報告書を引用し、これらの方法論に従ったと記載されている。したがって、この報告書における因果性の段階的判定も、これら二つの報告書の場合と同じように、偶然・バイアス・交絡という競合的解釈の影響を吟味しながら、それぞれの判定に割り振っているものと思われる。

その一方で、報告書は、四段階の段階評価法を提示した後で、Hillの9つの判定規準のそれぞれについて、詳細な検討を加えている。その際、Hillの判定規準が、「競合する非因果的な解釈に反対する根拠」(evidence against competing noncausal explanations) (いわば反証主義的)と「因果的な解釈を支持する根拠」(evidence supporting causal ones [explanations]) (いわば帰納主義的)という、二つの目的に利用できることを述べている。さらに、Hillの「規準を満たす数が増えるほど、より説得力のある競合的解釈を提示することが、次第に困難になる」と述べている。

つまりこの報告書では、能動喫煙と疾病との因果関係を評価するにあたり、「競合的解釈の否定」という(反証主義的な)立場を基本とする、段階的な評価法を導入している。その一方で、Hill流の判定規準について論じながら、「非因果的な競合的解釈の否定」という(反証主義的な)活用と、「因果性の積極的な支持」という(帰納主義的な)活用という、二つの方向性があることを指摘している。すなわち、反証主義的な「段階的評価」の中に、帰納主義的な「判定規準」を統合する試みとして、本報告書における因果推論の方法論を理解することができる。

## 7) 因果関係の評価の現状

3)から6)で述べた、因果性に関する各種機関の判定規準には、おおむね次のような共通点がある。

- ① 専門家の審査を経て出版された研究論文を、主な資料として採用している。
- ② 因果関係が「ある」「ない」という二分法ではなく、「十分な知見がある」「限定的な知見がある」「不適切な知見がある」などの段階的な判定を採用している。
- ③ 疫学研究と動物実験の双方を利用するが、最終的な判定にあたっては、疫学研究の知見をより重視している。例えば、三者いずれの判定においても、人集団を対象と

する疫学研究で「十分な知見」が存在しない限り、動物実験の知見のみに基づいて、因果関係の存在を肯定する最高位の判定（国際がん研究機関におけるグループ1など）を適用することは、原則的にない。

- ④ 疫学研究の評価にあたり、偶然・バイアス・交絡という競合的解釈を排除し得ている程度によって、研究の質を判断している。すなわち、因果性以外の競合的解釈という「誤り」が研究からどれだけ排除されているかという反証主義的な立場から、疫学研究の妥当性が判断されている。喫煙の健康影響に関する米国公衆衛生局長官報告のように、Hill 流の判定規準を活用する場合であっても、それに先立って競合的解釈の吟味が行われている。
- ⑤ 研究の進展に合わせて、判定の見直しと更新が行われている。

## 8) 結論

疫学は、実際の間人集団を対象として、化学物質をはじめとする暴露要因と健康障害との関連を検討するための科学である。したがって、化学物質のヒト健康影響を評価する際には、疫学研究からの知見が重要な役割を果たす。

とはいえ、化学物質のヒト健康影響について、疫学研究の知見のみで確定的な結論を導くことはできない。疫学研究では、主として、“化学物質の暴露量が多い人達が、少ない人達に比べて、ある病気になる確率が高いか否か” についての検証を行う。細胞レベルや分子レベルにおける化学物質の作用機序を、直接検証するものではない。また、化学物質による小さな影響を検出したり、複数の化合物の相互作用を評価するためには、非常に大規模な疫学研究を行うことが通常必要になるが、そうした研究の実施が現実的に困難な場合も少なくない。そのため、化学物質の作用機序を明らかにし、複数の化合物の相互作用等を検証するためには、培養細胞や実験動物を用いた基礎研究を行うことが必要になる。

ただし、これらの基礎研究も、実際の間人集団が化学物質に暴露する状況とは異なる状況で検討が行われるため、その結果をただちに人間の暴露状況に外挿できるとは限らない。したがって、化学物質のヒト健康影響について、疫学研究の知見のみで確定的な結論を導くことはできないのと同じように、基礎研究の知見のみで確定的な結論を導くこともできない。

結局、化学物質をはじめとする暴露要因とヒトの健康障害との因果関係の有無を、議論の余地なく完全に証明することは、疫学研究と基礎研究のいずれでも不可能である。そのことを前提とした上で、不完全な実証研究のデータに基づいて、できるだけ誤りの少ない形で因果関係を評価し、具体的な対策に結びつけるために考案された方法論として、上に紹介した各種の判定規準を理解することができる。わが国において、内分泌系に作用することが試験管内実験などで示されている化学物質のヒト健康影響を評価し、その対策を検討する場合も、因果推論の方法に関する国際的な現状を十分に踏まえて議論することが重要である。

## [文献]

Hill AB. The environment and disease: association or causation? Proc Roy Soc Med 1965;58:295-300.

National Academy of Sciences. Veterans and Agent Orange, Update 2002, Executive Summary. Washington DC: National Academy Press, 2003

National Toxicology Program. Report on Carcinogens. Ninth edition. Research Triangle Park, NC: U.S. Department of Health and Human Services, Public Health Service, National Toxicology Program, 2000.

International Agency for Research on Cancer. Some Internally Deposited Radionuclides. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Vol. 78. Lyon: International Agency for Research on Cancer, 2001.

U.S. Department of Health and Human Services. The Health Consequences of Smoking: A Report of the Surgeon General. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, 2004.