

4、肥満度の測定法

運動をして肥満を改善していくには、自分で肥満度をモニターしていくことが必要です。どの指標にも長所と短所がありますが、自分で1つの指標を決めて変化をみるのであればどの指標でも問題ないと思われれます。簡便に身長と体重のバランスをみるには body mass index (BMI)が、体脂肪量を測定するには生体インピーダンス法 (BIA) が優れています。

1) 体格指数法

成人 $\text{body mass index (BMI)} = \text{体重 (kg)} / \text{身長}^2 \text{ (m)}$

丹治指数 = $(\text{胸囲} + \text{腹囲} - \text{身長}) / 2$

加藤指数 = $(\text{身長} - 50) / 2$

乳児 $\text{カウプ指数} = \text{体重 (g)} / \text{身長}^2 \text{ (cm)}$

児童 $\text{ローレル指数} = \text{体重 (kg)} \times 107 / \text{身長}^3 \text{ (cm)}$

2) BMI による肥満の判定

身長 (m) の2乗に体格指数 BMI の22を乗じた数値を標準体重とする。

$\text{標準体重} = \text{身長}^2 \text{ (m)} \times 22$

標準体重に基づいて、次のような式により肥満度を算出する。

$\text{肥満度 (\%)} = (\text{実測体重} - \text{標準体重}) \div \text{標準体重} \times 100$

たとえば身長 170cm で体重 80kg の人の肥満度を求めるには、まず前述の式によって標準体重を算出する。

$\text{標準体重} = 1.7 \times 1.7 \times 22 = 63.58 \text{ (kg)}$

これを用いて上の肥満度の式によって算出すると、肥満度は以下のようなになる。

$\text{肥満度} = (80 - 63.58) \div 63.58 \times 100 \approx +25.8\%$

また身長 170cm で体重が 55kg の人の肥満度は以下のようなになる。

$$\text{肥満度} = (55 - 63.58) \div 63.58 \times 100 \approx -13.5\%$$

3) 標準体重法

(1) 箕輪の標準体重表

群馬県下集団健診を参考値として、1962年に日本人成人男女別の標準体重を作成したもの。1985年に高身長群の標準体重を追加して、男性は155～195cm、女性は145～185cmの標準体重を1cmきざみに表示している。が、男性の基準が現在の栄養状態に比べ低過ぎるなどの理由から、最近ではあまり使われていない。

(2) 肥満とやせの判定表（厚生省）

国民栄養調査成績を基にして作成されている。単に現代人の平均値を示したものであり、医学的に望ましい体重という基準で作られたものではないといえる。他の指標の標準体重と大きな差はないが、年齢の上昇とともに増加することを認めたところにも問題がある。

(3) 松木の標準体重表

Walkerによるメトロポリタン生命保険会社の理想体重表を基に、日本人用の単位に直し、換算して使用。死亡率が最低の体重を理想体重としている。

(4) 明治生命・塚本の標準体重表

各身長に対してもっとも死亡率の低い体重を標準体重として求めたもの。当時の日本人の平均体重に比べると若干重い数値であった。30～69歳に適用される。時代とともに変化する指標であるので、時々修正が必要。

4) 体脂肪量測定法

(1) 体密度法（水中体重法）

脂肪の密度は小さいが、筋肉中のタンパク質や骨中のミネラルは脂肪に比べて密度が大きいという原理を基に身体密度を求める。

体脂肪量が多い人ほど身体密度は小さく、逆に体脂肪量は少ないのに筋肉量が多いという人ほど密度は大きくなることを利用して体脂肪率を測定する方法。

体脂肪率の測定法として最も信頼されていて、“gold standard”と呼ばれ、他の測定法の基準とされているが、大がかりで被験者の負担大であるため、体力のある若者以外の測定は困難。

(2)生体インピーダンス法 (BIA)

筋肉、骨、脂肪の電気伝導度の違いを利用して測定する。除脂肪組織の約7割は電解質を多く含む体水分であり電流が流れやすく、脂肪組織や骨は電解質をほとんど含まない絶縁体であるため電流がほとんど流れない。生体に高い周波数の電流を流すと、電流は細胞膜を通して細胞内液にも流れる。電気インピーダンス（抵抗）はからだの電解質組成を反映するので、身長、体重から体脂肪率が推定できる。からだは測定原理のように理想的なモデルではないので、水中体重法から求めた補正式を使って体脂肪率を求める。

測定方法は、被験者の手の甲と足の甲に電極を置き、電流を通じ、同時に同側の手首と足首においた電極間の電圧からインピーダンスを計測する。最近では、電極を手足に装着する代わりに両足底の前方部から電流を通じて両足踵間の電圧から、インピーダンスを求める簡便法も開発され、体脂肪計として市販されている。

測定は非常に簡単で、安全、迅速、便利であることから、急速に普及している。注意点としては、

1. 測定時の体位を一定にする。
2. 皮膚と電極接触面をしっかりと密着させる。

3. 体温上昇を伴う運動などの後には測定しない。
4. 多量のアルコール摂取や発汗による脱水状態、または異常な摂水など、体水分の状態が正常でない場合の測定をひかえる。

また、腕で測る形式のものなどでは電極を常に一定の位置にしなければならない。生体インピーダンス法は体内の水分バランスの影響を強く受けるので数値の変動が大きいことや、電流が流れるのはからだの一部であるとの認識が必要。また、体密度法も可能性がある。

(3)近赤外線法

もともとは主に食品の主成分の分析に使われていたのを人体に応用したもの。生体に近赤外線の特長波長帯を照射すると、特異的に体脂肪に吸収され、一部は反射して戻ってくる。反射エネルギーは局所の脂肪量によって異なるので、これを測定すれば体脂肪量が推定できる。測定が簡単で使いよいが、肥満の脂肪量を過小評価し、やせの脂肪量を過大評価する傾向がある。

(4)皮脂厚測定法

皮下脂肪を測定し、肥満度を判定する方法。二点間をはさみ、その厚さを測定する器具を皮脂厚計という。肩甲骨下部と上腕部の2カ所について測定し、一般的にはその数値を足したもののから脂肪率を算出する方法が用いられる。合計した数値が、男性で40mm、女性で45mmを超えると、それぞれ体脂肪率が25%、30%のレベルを超えた肥満と判定される。測定方法がとても簡単なので、スポーツクラブなどでよく利用されているが、測定時の姿勢や皮膚のはさみ方、はさむ方向などによって誤差が生じやすい。

中高年労働者の食行動と血清レプチンレベル

分担研究者 畝 博 (福岡大学医学部衛生学)

研究協力者 百瀬 義人 (福岡大学医学部衛生学)

1、研究目的

脂肪組織から見つかったホルモンであるレプチンは、脂肪の貯えに関係する中枢神経系統へのシグナルであり、食欲をコントロールする摂食抑制作用がある。しかし、このレプチンと肥満との関係は複雑で、未だ謎に包まれたままである。レプチンが世界中から注目を集めているのは、体重増加の指標になり、肥満と関連する疾病の発症の指標になると期待されるからである。

ヒトにおいては、単回の摂食では血中レプチンレベルに有意な変化はみられないが、肥満者の減量療法の過程においては、日単位で血中レプチンレベルの低下が観察されている。このことは、レプチンが食行動と関連を持つ可能性を示唆している。一方、3年間の追跡研究によると、肥満進行群の血中レプチンレベルが、肥満が進行しなかった群と比較して低値を示している。このことは、レプチンの相対的な不足が体重増加につながる可能性を示すと考えられる。

レプチンは体脂肪率やBMI (body mass index)と正相関を示す一方で、個体差が大きいという特徴もある。したがって、同じ肥満度に対してレプチンレベルが明らかに高いグループ、明らかに低いグループを比較することにより、食行動の特徴の把握、さらには肥満発症リスクの予測に役立てることができると考える。本研究は、レプチンと食行動との関連の有無を検討し、さらに、レプチンレベルの低い肥満者の食行動の特徴を明らかにする。

2、研究方法

1) 対象

2000年度の定期健康診断を受けた、某企業労働者915名(40-65歳)のうち、現病歴(高血圧、糖尿病、心臓病、高脂血症)があつて服薬中の87名と、病気による禁煙者3名、TG \geq 1,000mg/dLの7名を除く818名を対象とした。

2) 方法

(1) 血液生化学検査

採血は早朝空腹時に行われた。レプチンならびに他の血液生化学検査値(血清脂質、HbA1c等)の測定は専門機関に依頼した。レプチンの測定はRIA法(radioimmunoassay)が用いられた。

(2) 食行動

坂田らが作成した食行動質問票(50項目)を用いた。この質問票の回答方法は、例えば、「自分は他人より太りやすい体質だと思うか」という質問に対し、「まったくその通り」、「そういう傾向がある」、「時々そういうことがある」、「そんなことはない」という、4つの選択肢の中から1つ選ぶ。選択肢に応じて、それぞれ4点、3点2点、1点が与えられており、これに基づいて7つのカテゴリー別(体質に関する認識、空腹感・食動機、代理摂食、満腹感覚、食べ方、食事内容、リズム異常)にスコア化した。

(3) 他の関連要因

BMI (body mass index; kg/m²)の算出には、健診時の身長と体重を用いた。血圧は水銀血圧計を用いて椅座位にて測定した。女性は閉経の有無を尋ね、閉経前後で2群に分けた。

(4) 統計解析

レプチンレベルは正規性を示さず、左に歪んでいたため、各要因との関連を検討する際は対数変換後の値を用いた。食行動とレプチンの独立した関連の検討では、共分散分析法を用いた。低レプチンレベルの肥満者における食行動の特徴の検討では、 χ^2 検定を用いた。すべてのデータ解析はSAS (statistical analysis system)を用いて行われた。

3、研究結果

1) 血清レプチンレベル

男性の平均値 (標準偏差) は 2.93ng/mL (± 1.47)を示した。女性は、閉経前群が 5.68ng/mL (± 3.10)、閉経後群が 5.67ng/mL (± 2.78)と同様のレベルを示し、閉経の有無で差は認められなかった。男女差を比較すると、女性の平均値が男性の約1.9倍高く、性差が認められた。

2) BMI とレプチンの関連

両者の相関関係を検討した結果、有意な正の相関関係が男女ともに認められた (男性: $r=0.476$, $p<0.001$, 閉経前女性: $r=0.560$, $p<0.001$, 閉経後女性: $r=0.512$, $p<0.001$)。血圧、血清脂質、HbA1c による補正後においても、BMI とレプチンとの関連の有意性が認められた (表1)。したがって、BMI が高いほどレプチンレベルも高いことが明らかになった。

表1. BMI 別にみたレプチンレベル (補正平均値)

BMI	男性			閉経前女性			閉経後女性		
	n	M	SE	n	M	SE	n	M	SE
やせ	59	0.274 (0.038)		27	0.591 (0.072)		14	0.556 (0.061)	
普通	325	0.395 (0.029)		80	0.762 (0.059)		39	0.668 (0.036)	
過体重	145	0.493 (0.031)		24	0.926 (0.069)		7	0.924 (0.086)	
肥満	82	0.546 (0.033)		10	1.075 (0.085)		6	0.897 (0.093)	
p for trend		$p<0.001$			$p<0.001$			$p<0.01$	

n:人数, M:補正平均値, SE:標準誤差

BMI; やせ : <19.8, 普通 : 19.8-24.1, 過体重 : 24.2-26.3, 肥満 : ≥ 26.4
 補正平均値 : 血圧、血清脂質、HbA1c により補正を行なった。

3) 食行動とレプチンの関連

BMI で補正した結果、男性にのみ 2 項目 (体質に関する認識、満腹感) で有意な正の関連が認められた (表 2)。すなわち、この 2 つのカテゴリーに属する食行動がレプチンレベルを高めている可能性を示唆した。

表 2. 男性における、食行動スコア別にみたレプチンレベル (補正平均値)

	体質に関する認識			満腹感		
	n	M	SE	n	M	SE
低スコア群	285	0.384 (0.013)		293	0.394 (0.012)	
高スコア群	326	0.450 (0.012)		318	0.444 (0.012)	
p 値	p<0.001			p<0.01		

n:人数, M:補正平均値, SE:標準誤差
 補正平均値 : BMI により補正を行なった。
 食行動スコアは中央値で 2 群に分けた。

4) 低レプチンレベルの男性肥満者における食行動の特徴

表 3 に示したように、低レプチンレベルの肥満者は 36 名であり、全男性の 5.9%、肥満男性 (82 名) の 44% を占めた。

表 3 低レプチンレベルの男性肥満者における食行動の特徴

	体質に関する認識			合 計
	低スコア群	高スコア群		
低レプチン (<2.6ng/mL)				
かつ肥満 (BMI ≥ 25) の群	2 (5.6)	34 (94.4)	36	(100.0)
上記以外の群	283 (49.2)	292 (50.8)	575	(100.0)
合計	285 (46.6)	326 (53.4)	611	(100.0)

Manter-Haenszel Chi-square DF=1, Value=25.9, p=0.001.

彼らの食行動の特徴を他の男性群と比較した結果、「体質に関する認識」が高かった ($p=0.001$)。しかし、「満腹感」においては、有意差は認められなかった ($p=0.437$)。

4、考察

1) 血清レプチンレベル

女性のレプチンレベルが高く性差が認められたこと、閉経の有無に影響されていなかったことは、他の報告でも示されている。性差について検討した研究では、テストステロンとレプチンレベルとの間に負の関連が見られている。テストステロンの抑制的影響が1つの要因となって、男性のレプチンレベルが低くなっていると推察される。

2) BMI とレプチンの関連

正の独立した関連が認められたことは、他の報告と一致している。BMI は必ずしも体脂肪量を反映しておらず、異なる結果をもたらす可能性があるが、多くの報告においてレプチンレベルと関連が認められたことは注目に値する。

レプチンレベルが上昇しているにもかかわらず、BMI が高値で肥満が是正されていない理由は、肥満形成におけるレプチンの作用機構の障害が関与していると考えられている。基本的にレプチンは、脂肪細胞で生合成されたのち循環血中に分泌され、視床下部のレプチン受容体に直接作用することにより、強力な摂食抑制を促し、体脂肪量を一定に保つフィードバックループを構成している。このことから、レプチンレベルが高いにもかかわらずBMI が高値を示すのは、個人のレプチン受容体の数が少ないのかも知れないし、信号経路に対する反応が低下しているのかも知れない。あるいは、レプチンによる血液と脳との間の連絡が乏しいのかも知れないし、レプチンの活動が少ないのかも知れない。したがって、レプチン受容体の構造、あるいは伝達経路の解明が重要と考えられる。

3) 食行動とレプチンの関連

男性において、「体質に関する認識」と「満腹感覚」がレプチンレベルのコントロール要因である可能性がうかがえた。すなわち、「自分は他人より太りやすい体質だ」と考えている人ほどレプチンレベルが高いということから、「体質の認識」と行動のズレが考えられる。肥満は遺伝と関連し、個人差は確かに認められる。しかし、肥満の原因を自分の体質と決めつけて、何の努力もしないことが肥満につながり、結果としてレプチンレベルを高めることになる。「満腹感覚」のスコアが高いということは、過食につながる。食後でも好きなものなら食べてしまうし、食べ過ぎを他人からよく注意される。料理が余るともったいないので食べてしまうし、たくさん食べたことを後悔する。

4) 低レプチンレベルの男性肥満者における食行動の特徴

「体質に関する認識」が高かった。低レプチンレベル、すなわち、レプチンの不足は摂食抑制作用の低下を意味しており、肥満につながる。彼らの食行動の特徴として「体質に関する認識」のみが高かったことから、レプチンの不足が「太りやすい体質」として認識されているのかも知れない。

低レプチンレベルの肥満男性は 36 名で、全肥満男性（82 名）の 44%を占めた。残りの 56%は高レプチンレベルの肥満者であり、レプチンが分泌されても効果がみられない群と言える。その原因を明らかにするとともに、フォローアップ・スタディによる体重の増加とレプチンレベルとの関連を検討する必要性がある。

5、結論

中高年労働者（年齢 40～65 歳）の健康評価の情報源とするため、特にレプチンと食行動に注目した。このレプチンに対する性、BMI ならびに食行動との関連を検討し、次のことが明らかになった。

1. レプチンレベルは性差が認められ、女性が男性の約 1.9 倍高かった。

2. BMIとレプチンレベルとの有意な正の関連が、男女ともに認められた。

3. 食行動とレプチンの関連を検討した結果、男性において「体質に関する認識」と「満腹感覚」で有意差が認められた。これら2項目の食行動がレプチンレベルを高めている可能性が示唆された。

4. 低レプチンレベル (<2.6ng/mL)の男性肥満者 (BMI \geq 25)における食行動の特徴は、「体質に関する認識」が高かった。レプチンの不足が「太りやすい体質」として認識されているのかも知れない。

レプチンレベルが体重増加の指標になるかを明らかにするためには、食行動調査を含めて体重増加とレプチンレベルとの関連を検討する必要がある。

6、研究要旨

中高年労働者（40～65歳の男女818名）のレプチンと食行動との関連の有無、並びにレプチンの低い肥満者の食行動の特徴について検討した。レプチンは男女差を認め（男2.9ng/mL,女5.7ng/mL）、BMI(kg/m²)とは正の関連を認めた。閉経の有無はレプチンと関連しなかった。食行動とレプチンとの関連を検討した結果、男性における「体質に関する認識」と「満腹感覚」で有意な正の関連を認め、これらの食行動がレプチンレベルを高めている可能性が示唆された。低レプチン(<2.6ng/mL)の男性肥満者(BMI \geq 25)は、「体質に関する認識」が高かった。レプチンの不足が「太りやすい体質」として認識されているのかも知れない。BMIと食行動はレプチンレベルの有意な決定要因であった。これらの結果は、BMIと同様に食行動が中高年労働者のレプチン分泌に対するコントロール要因であることを示唆している。

現実の介入研究や臨床試験において生じる 好ましくない事態への対応に関する考察

分担研究者 津田 敏秀 (岡山大学医学部衛生学教室)

1. はじめに

介入研究や臨床試験は、ランダム割付が行われている場合にはランダム化臨床試験(RCT)として、その研究による結果は最も信頼に足るものとされると信じられている。これに加えて、二重盲検法が施されている場合には、さらに信頼に足るものとされ、疫学研究における Gold Standard とされることが多い。しかし、現実にもそのような研究を実施することは、様々な困難があることもまた指摘されているところである。また、実際に実施が困難にも関わらず、RCTの目的は、未知の交絡要因を調整しているはずであるということしかない。従って、実施上のコストや達成の困難さが、RCTのメリットを上回るのであれば、実施上のメリットはなくなることになる。従って、RCTが実際に実施されるときに予想される欠点を理論的に押さえておく必要がある。

まず、観察研究に比べ、介入研究はそれ自体が、不自然な状況を作り出しており、そこから得られる結果を日常の診療に一般化できるかどうかの問題になる。次に、二重盲検法は、医師にも患者にも気づかれる場合があり、その場合にはバイアスが入ることが指摘されている(Norell 1995)。例えば患者が振り分けられた治療に気づいた時には、別の治療に走ることになる。この時は交絡バイアスが入ることになる。また、追跡不能例が生じたり、コンプライアンスがなかったりすることが、片側グループだけに偏って生じた時に選択バイアスが生じることになる。また、医師が患者の割付に気づいた時には、differentialな疾病の誤分類を導入する可能性が生じる。これらによって二重盲検法は成り立たなくなる可能性がある。さらに、介入自体は誰が見ても明らかなものであり、どちらの群に属したかが誰の目にも明らかな時には、二重盲検法は機能しない。

次にランダム割付の役割について述べる。ランダム割付が観察研究に比較して有利な点は、未知の交絡要因を調整することである。しかし、すでに述べたように観察研究に比較した欠点は、そこから得られる結果を日常の診療に一般化できるかどうかの問題である。

本研究が示す知見の意味を十分に理解し、日常の保健活動に本研究の知見を役立てていただけるために、まず、本研究のいくつかの特徴を、短所も含めて、簡潔にまとめて明示する。続いて、疫学研究における誤差およびバイアスについて概説し、本研究の特徴を考察する際の必要な基礎的知識を示す。その上で、本研究の特徴のそれぞれについて、例示しながら考察を進める。最後に「まとめ」として、医学・医療研究における疫学研究の意義を示すと共に、本研究の特徴に関して簡潔にまとめる。

2、本研究の特徴

以上の点を踏まえて、本研究デザインにおける問題点となりうる特徴を列挙してみる。

1) まず、昔の介入群に対しそれに相当する対照群を作り、ベースラインから毎年検診で両群を比較する。具体的には、ベースラインもしくは介入群に対して過去についての情報を集めることになる。つまり、毎年検診をしている集団があり 1995 年にその集団の中における肥満者の一部がヘルスセミナーを受けたとする。1994 年にさかのぼって、ヘルスセミナーを受けた介入群に対して、対照群を設定し、(肥満集団)1996 年、1997 年、1998 年を比較していこうという考え方をおこなっている点である。いわば、ヒストリカル介入とでも言うような状態である。

2) 次に、毎年、ヒストリカルな介入前の検診と介入後の検診で両群を比較することになるが、結果が年ごとに異なる可能性がある点をどのように処理するかという点である。

3) 最後に、無作為比較対照ができた年もあるが、その時、比較群に比較して介入群において参加者が少なくなる可能性がある点である。

以上の三点が、本研究の長所であり欠点でもあり、そして研究結果の現実である。「はじめに」において述べたように、一概に介入試験・臨床試験とはいっても、様々な比較群(非介入群)の長所と欠点がある。症例対照研究の対照群の選択方法にも様々な方法があり、それぞれの長所と欠点があるように、臨床試験においても、現実的には上記のような様々な比較群の選び方があり、それぞれの比較群において長所と欠点がある。従って、以下においては、本研究の比較群の選択方法に潜在する可能性がある複数の比較群の選びかたを中心に、様々な比較群の選び方のそれぞれについて、長所と欠点を考察し、その結果に及ぼす影響を予想する。

なお本研究は、結果判定が医師の診察結果ではなく、検査値によって比較がなされている。従って、「はじめに」で述べたような differential な疾病の誤分類は生じ方は少ないと予想される。

3、疫学研究における誤差およびバイアスについて

1) 偶然による誤差(チャンス)

いかなる測定誤差も、偶然による変動が含まれている。しかし、この変動は、測定回数が十分に大きい時には、平均するとゼロになることが期待できる。従って、試行回数が大きくなると結果として測定の誤差は小さくなり、その結果、測定の精度は増加すると考えられる。測定回数(標本数)の大きさを考慮して偶然による誤差(チャンス)を評価する推定値は、点推定値に対して区間推定値と呼ばれる。通常、95%信頼区間が選択されることが多い。なお点推定値は 0% 信頼区間である。そして 100%信頼区間とは推定値が取り得る値の全ての範囲ということになる。

検出したい介入群と非介入群の影響の異なりに比較して標本数が少ない場合には、検出力不足(lack of power)の問題となる。従って、有意差がないからといって、影響の違いがないと判断した場合、本来は(標本数が多い場合)検出されるべき違いが検出されないという事態が生じることになる。この、いわゆる

「偽陰性」の状態を生じることに関しては、観察データを見る場合には注意を払うべきであろう。

2) 交絡バイアスと反現実モデル(counterfactual model)

疫学研究における仮説は、ある曝露(もしくは介入、外的要因)が疾病発生の頻度に影響するかどうかということで設定されることが多いが、現実においては、曝露は仮説として挙げられる要因のみだけではない。そのような疾病発生の頻度に影響し、測定値としての疫学的影響の指標に変動(バイアス)を及ぼす可能性がある要因のことを交絡要因候補と呼ぶ(potential confounder: 「共変量」と呼ぶこともある)。このような交絡要因候補は仮説の一部としてデータが集められ、分析の際にモデルの中に組み込まれて推定の際に調整されることが多い。

交絡要因であるためには次の三条件を満たしておかなければならない。このうち第三の条件は省略されることがある。

- (1) 交絡要因は当該疾病のリスク要因でなければならない。
- (2) 交絡要因は、当該曝露に関して人口母集団において関連していなければならない。
- (3) 交絡要因は、当該曝露と当該疾病の因果の途上にあってはならない。

このような条件を満たして、初めて交絡要因候補は交絡要因として推定値にバイアスを生じさせ得る。

これまで、交絡要因候補が実際に交絡要因として交絡バイアスをもたらしたかどうかについては、change in estimate として、当該交絡要因候補をモデルの中に組み込んで調整した(adjusted)推定値と、モデルの中に組み込まずに推定したとき(crude)の推定値とが異なるかどうかで判断されていた。しかし、これでは交絡するか否かが、どのような分析モデルを選択したかによって交絡があつたりなかつたりすることになる。例えば、推定値を率差とするか率比と

するかによっても、交絡要因であるか否かが異なってくる。

このような欠点を回避するため、曝露した集団が、「もし曝露していなければ・・・」というような反現実的な状況(counterfactual model: 反現実モデル: 後注)を基準として、現実の非曝露集団と比較した場合、非曝露集団のずれを交絡と定義する場合もある(Rothman 1998)。この定義の長所は、選択した分析モデルに交絡の有無が影響されないということである。一方、単なる理想的な対照群というような目標を設定するだけとは言え、現実にはあり得ないモデルを設定し、それに言及すること自体は欠点とも言える。

3) counterfactual model (反現実モデル)

1920年代に薬学分野で唱えられ初め、1950年代頃より哲学の分野で語られ、最近はや学で盛んに説明に用いられている反現実モデル(counterfactual model)を、Pearlは因果論のテキスト(Pearl 2000a)の中で次のように説明している。

「例えば、治療Xが対象者の反応Yの分布(これは回復 $Y=0$ もしくは死亡 $Y=1$ をとる)に何らの影響もない、ある臨床試験(例えばランダム化された)において、ある対象者、Joeを仮定した時、彼が治療を受け、そして死んだと仮定しよう。即ち、我々はJoeの死が、治療のために、あるいは治療にも関わらずもしくは治療に関係なく、生じたかどうかを問うのである。他の言葉で言うと、我々はJoeがもし治療を受けなかったら彼が死んでいたであろう確率Qを問いかけるのである。

そのような反事実的問いに答える困難さを際立たせるために、治療群と比較群の両方で患者の50%が回復し50%が死亡したという極端な例を考えてみよう。さらにサンプルサイズが無限大に近づいたと仮定する。従って、

全てのxとyに関して、 $P(y|x) = 1/2$

統計学的検定に通じている読者は、利用できるデータからは、この反事実的問いに答えることは不可能であることを即座に認識するであろう。すなわち治療を受けて死んだJoeは、決して治療を受けない状態では検証されないのであ

る。さらに、この困難さは Joe という特定の個人への問い、つまりある 1 点のデータからのみ得たという設定をしたことから生じるのではない。治療を受けて死んだ対象者群でどれくらいの割合 Q で、もし彼が治療を受けなければ回復したであろうか、と問う集団における頻度に関する問題として言い換えることは、これらの対象者群の誰もが非治療的条件の下では検証されないという理由で同じ困難性に遭遇するであろう。そのような困難さゆえに、形而上学的問いとして、反事実的問いから統計学者達は答えることを避けてきた。そして、直接的検証によって答えられることが可能な問いにのみ、統計学的分析を限ると唱えた。

しかし、反事実的言辞で苦しめられているのは、我々の科学的、法的、そして通常の言語であり、これらははっきりと反事実が形而上であってはならないことを示している。即ち、それらは明確に検証可能な運用をしなければならぬし、大事な本質的な情報をもたらさなければならぬ。」と Pearl は述べ、さらに「『反事実』という単語は不適切な誤称である。なぜならそれは事実を反した立場に依拠した言明を内包している、もしくは、少なくとも経験的な証明から逃げている言明である。反事実というのは、次のいずれの範疇にも属さない。即ち、科学的思考の基本的なものでも、何らかの科学的法則として経験的なメッセージを明白なものをもたらすものでもない、」と述べている。不可能なことを実際に語るにより科学的法則が語れるのだと述べ、通常語られる科学法則に関して反事実的表現を用いた時にいったいどうなるかという例を幾つか挙げて、それが通常語られる科学法則の言い方とは異なっていることを明示している(Pearl 2000b)。

4) 情報バイアス

情報バイアスは、曝露の測定誤差(誤分類)、疾病の測定誤差(誤分類)などが生じることにより、疫学的影響の指標の推定値に生じるバイアスのことを言う。

ある研究対象集団において、対象は調査され、曝露と疾病に関して分類される。曝露の誤分類 *exposure misclassification* は、曝露された人が誤って非曝

露と分類された時、もしくはその逆の時に生じる。同様に、疾病の誤分類disease misclassificationは、病気の人が誤って病気でないと分類された時、もしくはその逆の時に生じる(Norell 1995)。

病気の発症・有無に関わりなく、曝露の誤分類が生じた時（感度もしくは特異度が1より小さい時）には、率比等の相対危険度、すなわち影響の程度の指標は一部の例を除いて影響が相対危険度1(RR=1)の方向にバイアスされ（non-differential exposure misclassification）、曝露の有無に関わりなく、疾病の誤分類が生じた時（特異度が1より小さい時）には、影響の程度の指標はRR=1の方向に評価されるかunbiased（特異度が1のとき時）となる（non-differential disease misclassification）。

Differential exposure misclassificationで最も有名な例としてrecall biasがある。recall biasは、これまでのデータから判断してそんなに頻繁に生じるものではないが、症例が生じた後に曝露の情報を収集する研究では可能性があることと認識するべきである。例えば、妊娠中の薬剤服用による奇形児出産に関する研究で、症例群の曝露歴に関して奇形児を出産した母親から聞き取り、対照群の曝露歴に関して健常児を出産した母親から聞き取ったとする。この時、奇形児を出産した母親は健常児を出産した母親に比較して、薬の服用歴を詳しく思い出そうとする傾向があることが予想される。この場合、影響の程度の指標は過大評価される方向にバイアスされる。疾患罹患の有無により、曝露の誤分類の程度に違いが生じたからである。このようなバイアスに関しては、母親から聞き取った曝露歴による判断とカルテに記載された曝露歴による判断を比較することにより、ある程度、評価することが可能である。実際、この2つの曝露評価の間には、ほとんど異なりがないか、あるいはあったとしてもわずかであることが知られている(Norell 1995)。

これに対し、non-differential exposure misclassificationは、疾患の有無により曝露の誤分類が影響されないことを意味する。例えば上記の例において、曝露歴の情報を妊娠中のカルテから収集したとしよう。このときカルテの記載漏れとか、担当医師の処方以外の薬を母親が服用するとかにより曝露の誤分類

が生じうる。しかしこの曝露の誤分類は、疾患の有無により影響を受けていないのでnon-differential exposure misclassificationと言える。この時、影響の程度の指標は相対危険度1の方向にバイアスを生じる。なおnon-differential misclassificationにより影響の程度の指標がどのように相対危険度1の方向にバイアスされてしまうかについては、Rothmanがテキストの中で飲酒と喉頭癌の関係の例を用いたシュミレーションの結果を提示している(Rothman 1986)。

一方、differential disease misclassification は、曝露の状態によって疾病の誤分類の程度に差が生じてしまうバイアスであり、non-differential disease misclassificationは、曝露の状態によっては影響を受けない疾病の誤分類により生じるバイアスである。differential disease misclassificationの例としては、喫煙と肺気腫の関係の研究で肺気腫の診断を胸部X線像で行う時の例が挙げられる。この時、医師が読影を本人と面談しながらもしくはカルテを見ながら行うと、喫煙者において肺気腫の診断を下しやすくなる現象が予想され、その場合、率比を過大評価する方向のバイアスが招かれることが予想できる。一方、胸部X線写真だけで独立して放射線科医が読影を行うと誤分類はかなり生じるであろうが、これはnon-differentialな誤差を生じるであろう。

疫学ではnon-differentialな誤分類の方が,differentialな誤分類よりもバイアスの方向性がはっきりしているので、前者よりも後者の誤分類を避けようとする。non-differential misclassification, differential misclassificationの位置づけは、情報バイアスinformation biasとして通常は分類される。

5) 選択バイアス

選択バイアスは、研究対象を選択する過程で生じてくるバイアスである。研究の目標である理論的な研究対象(source population)での影響の推定値と、研究で選択された対象からの影響の推定値が異なることである。健康労働者効果(healthy worker's effect)や症例対照研究における対照の選択によるバイアスなどの例がある(referent selection bias)。状況により、影響の指標を過大評価する場合と過小評価する場合がある。

ランダムに割り付けた介入研究における選択バイアスは、選択的追跡不能例(selective loss to follow-up)の出現である。すなわち、介入群(曝露群)において、非介入群(非曝露群)と比較して、選択的に(曝露に関して)追跡不能例が多い場合と、少ない場合が生じる場合である。前者の場合は、介入による影響の指標を過小評価するし、後者の場合は、介入による影響の評価を過大評価することになる。

従って、介入群において曝露に関して選択的追跡不能例が生じた時にさえ、介入による影響の指標が増加していることが観察されたときには、もし介入群において追跡不能例がない場合には、実際の介入による影響は観察結果よりずっと大きいことが予想されることになる。

このように、研究毎に研究デザインや介入の性質に依存してバイアス結果は異なるので、研究にバイアスが介在しているからという理由で研究結果を全部否定することを行うようなことをせずに、介入の性質やバイアスの方向性を検討し考察の上、妥当な結論や今後の研究課題を提示することが大事であろう。

6) 因果関係について－Hume が科学研究に与えた科学哲学的意義と本研究

先行する事象(event)とそれに引き続いて後で生じた事象(event)との両方の事象が生じたことは、合意に達することが可能であるが、先行する事象と後で生じた事象との間に、前者を「原因」とし、後者を「結果」とする因果関係が存在するか否かに関しては、合意に達することは不可能である。これが、Hume が 18 世紀に世に問いかけた因果関係に関する懐疑論である。因果関係なるものが存在することによって、日常生活が成り立っていることを考えると、この根本的な懐疑論に対して説明もしくは反論を加えられないことは、哲学や後に隆盛を迎えることになる科学の世界に衝撃を与えた。

疫学における因果推論の議論が、Hume の問題から開始されることが多いのは、疫学自体が、先行する曝露とこれに続く疾病の発生の因果関係に関する問題を取り扱っているからであろう。疫学という方法論の発達は、この Hume