

# Prevention Trials of Long Duration

しばしば非常に長期に渡る因子への暴露が病気につながります。喫煙と肺癌のように、10 倍以上にリスクを押し上げるような場合には、どのような臨床研究でも正しい結果を導きやすいと思います。しかしながら、20-30%程度(1.2-1.3 倍)病気になる確率を押し上げるかどうかを検討するには、統計学的パワーを必要とするため、何千人という多くの人々の参加を必要とします。今までもフレミングハートスタデーをはじめとする大きな prospective cohort study が行なわれ重要な発見がありました。しかし、観察的臨床研究において confounder の存在をどうしても除外することができない弱点があります。特に生活習慣と慢性疾患の発生に関しては prospective cohort study では confounder が関与しやすいのが現状です。例えばビタミン摂取が発癌を抑制したとしても、元来健康に気を使う人はビタミンだけでなくジョギングをしたり、ダイエットをしたり、健康によいとされるありとあらゆるものを励行する傾向があり、どれが大切な要素であったかわかりません。そこでこれらの confounders を除外するため、生活習慣病に対しても近年大規模な Randomized trial が行なわれるようになりました。

ここで挙げた臨床試験は beta-Carotene and Retinol efficacy trial (CARET)、すなわちベータカロチンが肺癌発生を減少させるかを検証しようとした、大きな臨床試験の 1 つです。予想に反して、ベータカロチン群で肺癌の発生が多い傾向が確認され、当初の予定より早期に研究を中止しています。

臨床試験では倫理的問題が絡んでくる為、観察的研究と異なり、途中で中止せざるを得ない状況も発生します。どのような状況で臨床試験を中止するべきなのでしょうか？

# Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease.

Omenn GS, et al. New Eng J Med 334:1150-5,1996.

## 要旨

肺癌は米国において最も重要な死因の1つです。Carotenoids や retinoids による肺癌予防効果の可能性が示されています。そこで我々は多施設共同ランダム化プラシーボ併用二重盲検臨床試験を行ないました。この Beta Carotene and Retinol Efficacy Trial には 18,314 人の喫煙者（過去喫煙者あるいはアスベスト暴露者）に対して 30mg のベータカロチンと 25,000IU のレチノールの合剤を用い、肺癌発生の incidence を経時的に検討しました。73,135 person-years（平均4年の観察）で 388 人の新しい肺癌患者が発生しました。治療群はプラシーボ群に対して 1.2 倍( $p=0.02$ )の肺癌発生をみました。また全ての原因を含む死亡は治療群で 1.17 倍、肺癌による死亡に関しては 1.46 倍に治療群で有意に上昇していました。これらの結果より本臨床試験は予定より 21 ヶ月早期に中止となりました。今後5年間更に経過観察を継続する予定です。

## Study Questions

1. この臨床試験における null hypothesis は何ですか？ null hypothesis を最も強く規定する暴露、一過性の因子、結果の定義について考えてみてください。
2. 本試験が最も影響を受けるもっともらしい alternative hypothesis の特徴は何ですか？ 暴露の程度、時間因子、疾患修飾因子、参加者の性格について考えてみてください。
3. confounding で肺癌と治療薬との間に観察された関係を説明できますか？
4. 2つの肺癌になりやすい集団が選ばれました。臨床試験において2つの集団を含む利点と欠点は何でしょうか？
5. 治療群とプラシーボ群の間で肺癌の診断に違いがあった可能性はありますか？ ベータカロチンとレチノールをよく使用する population で調査できると仮定して、純粋な観察研究と介入研究とではどのような違いがありますか？
6. どの期間、肺癌の発生頻度は治療群よりプラシーボ群で高かったですか？
7. 途中解析によって介入研究を中止することがあります。観察研究でも同様のことがありえますか？

## 方法

### study design

簡単に述べると、CARET は 1983 年に組織され、1985 年 2 つのパイロット研究としてランダム化がはじまりました。1 つはかつて仕事上アスベストに暴露されたことのある 816 人の男性を治療群(15mg beta-carotene + 25,000IU retinol)とプラシーボ群に 1 : 1 の割合でランダムに振り分けました。もう 1 つは、ヘビースモーカーの男女 1029 人に対して同じように治療群(30mg beta-carotene + 25,000IU retinol)とプラシーボ群に 1 : 1 の割合でランダムに振り分けました。更に研究は参加施設を拡張して 1988 年から 1991 年まで延長されました。その際、ベータカロチンの量を 30mg に統一しています。1997 年末、110,000 person-years を目標としました。

### Eligibility, Recruitment, and Randomization

Randomization の 15 年前に最初にアスベストに暴露された、あるいは胸部レントゲンでアスベスト関連肺病変を認める 45-74 歳 (後 45 - 69 歳) の男性労働者、他アスベスト暴露の可能性が高い職種に 5 年以上就いた男性を対象としました。パイロット研究では喫煙の規定を設けませんでした。続く研究ではアスベスト以外に喫煙もしている人々を含めました。またもう 1 つのグループとして、マネージドケア組織より 50 - 69 歳の年間 20 箱以上のタバコを吸う喫煙者の男女を集めました。参加者はビタミン A 製剤の摂取を 5500IU 以下、ベータカロチン製剤の摂取を 0 にすることに同意しています。合計 4060 人のアスベスト労働者、14,254 人の喫煙者が randomization されました。1994 年 alpha-Tocopherol, beta carotene cancer prevention study (ATBC)の結果がでてからは、その結果の詳細を参加者に説明し、同意をとっています。

血液中のビタミン濃度を定期的 (2 年毎) に測定しました。死亡以外の理由で治療を途中で中止したものも最後まで経過観察し、解析に加えました (注 \* 本研究では服用の状況を客観的に測定できる点大きな強みとなっています。しかし半減期にもよりますが、血中濃度があっていたからといって毎日服用していた保証はありません。)

### データ回収、安全性モニター、End points

参加者は年に 1 回センターを受診してもらい、4 ヶ月に 2 回は電話連絡を受けています。97% は予定とおりコンタクトをとることができました。1995 年 12 月 15 日の時点で 98% の安否が把握されていました。臨床症状や新たに発生した病気に関してコンタクトをとる毎に質問用紙に従ってなされました。その際 13 の症状をモニターし、CARET symptom-assessment scale に従って点数化しました。またセンター受診の際には、診察に加えて肝機能等の血液検査も施行しました。

これとは独立した形で、safety and end-points monitoring committee が患者コードに基いた情報を用いて年に数回会合を開いていました。ATBC cancer prevention trialの結果が閲覧可能になった時点より、committee は blind system を解除しました (参加者は blind のまま) (注 \* 中止の判断には他の類似臨床研究にも注意を払う必要があります。)

### End points の評価

それぞれのセンターからの癌および死亡の報告は中央センターに集められました。臨床記録と病理標本は 2 人の腫瘍学者、2 人の内科医、1 人の病理学者により独立に (患者がどちらの治療を受けたか知らない状態で) 検討しなされています。1995 年 12 月 15 日、1446 人の癌と 974 人の死亡、合計 2420 の end points が確認されています。

## 結果

### 統計解析

肺癌の incidence に関してリスク（アスベスト対喫煙者）、時間（パイロット対その後）、6つのセンターに分けて、治療群とプラシーボ群間の差を weighted log-rank test で検討しました。Weight の有無は結果に影響なく、ここでは non-weight data をしめしてあります。Relative risk と confidence interval は stratified Cox regression model により算出されました。結局 67,449 person-years に対して 354 人の新たな肺癌と 829 人の死亡を確認しています。1994 年の時点で  $p < 0.0006$ 、1995 年の時点で  $p < 0.007$  を確認し、1995 年 12 月 15 日に試験を中止しています。

Table 1. Risk factors among the participants at base line  
Study population が多いため、分布に偏りはみられませんでした。

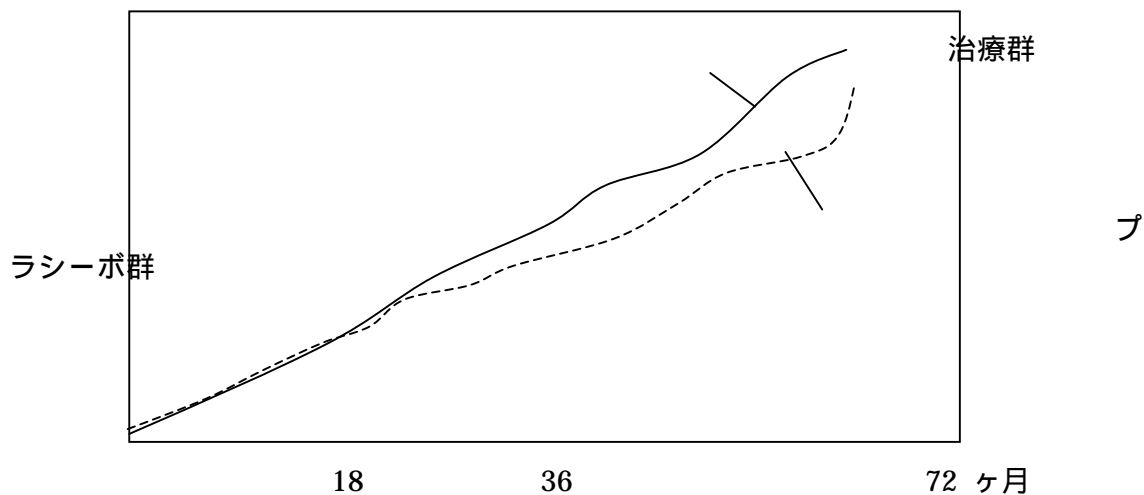
Table 2. Incidence of lung cancer in relation treatment with beta carotene and vitamin A

	No of subjects	Person-years	Case of lung cancer	Incidence (cases/1,000 PY)
Active treatment	9420	38,535	228	5.92
Placebo ratio	8894	34,600	160	4.62
				1.28
				$p=0.02$
95% CI				(1.04 – 1.579)

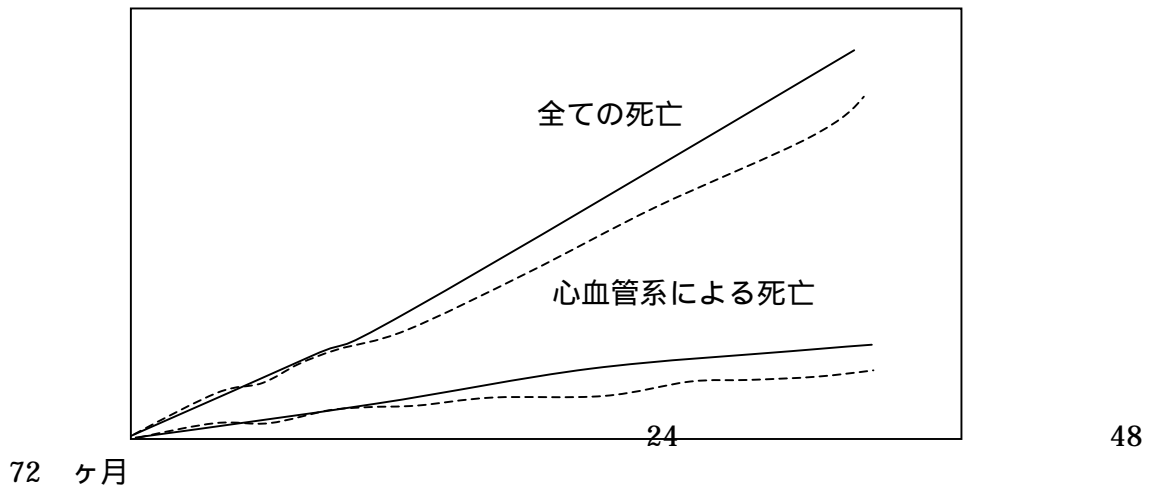
Figure 1. Kaplan-Meier curves of the cumulative incidence of lung cancer among participants receiving active treatment and those receiving placebo.

最初の 18 ヶ月において差は認められませんでした。それ以降差が開きました。

Cumulative incidence of lung cancer



Cumulative mortality



### 考察

著者らは「参加者は臨床試験に非常に協力的であった」と評価しています。また観察的疫学研究においてベータカロチンやビタミン A を多く含む野菜や果物の摂取が肺癌の incidence を抑制することが報告されているにもかかわらず、本研究においては治療群で 28% 高い肺癌 incidence を認めました。また全ての原因による死亡率、冠動脈疾患もそれぞれ 17%、26% 増加していました。これはフィンランドで行なわれた ATBC cancer prevention study と一致するものでした。Safety and end points monitoring committee と steering committee は治療群で肺癌をむしろ助長する結果となり、これ以上臨床試験を継続しても無駄であると判断し、中止を勧告しました。

我々は何故肺癌が治療群で増加したか説明することができません。

## Study question に対する解答

1. この臨床試験における null hypothesis は何ですか？ null hypothesis を最も強く規定する暴露、時間因子、結果の定義について考えてみてください。

Null hypothesis をアスベスト暴露者などに絞った場合、結論をアスベスト暴露者に対してしか適用することができず、一般に対して拡大することはできません(loss of generalizability)。これは bias とは異なったもので、external validity と表現されます。本研究における null hypothesis とは「6年間に渡るベータカロチンとレチノールの摂取はアスベストに暴露された中年男性と年間20箱以上の煙草を吸う男女の喫煙者に対して、肺癌の incidence に影響しない。」です。

2. 本試験が最も影響を受けるもっともらしい alternative hypothesis は何ですか？ alternative hypothesis とは「6年間に渡るベータカロチンとレチノールの摂取はアスベストに暴露された中年男性と年間20箱以上の煙草を吸う男女の喫煙者に対して、肺癌の incidence に影響する。」であり、null hypothesis の逆です。

3. confounding で肺癌と治療薬との間に観察された関係を説明できますか？

Confounding では説明が付きません。本臨床試験では非常に多い参加者を対象に行なわれ、表1をみてもわかる通り、2群間に confounder が存在する程の偏りを生じていません。また randomization を行なっているので、p値がどんなに小さくても、p値を base line measurement の表に示すことは意味がありません。

4. 2つの肺癌になりやすい集団が選ばれました。臨床試験において2つの集団を使用する利点と欠点は何でしょうか？

2つの肺癌になりやすい集団、すなわちアスベストに暴露された男性(その多くは喫煙者でもある)と、ヘビースモーカーの男女が対象として選ばれました。これら2種類の暴露因子は肺癌発生に関して異なった機序で働き、またそれぞれの cohort は別個に検討しても大丈夫な程の数を備えています。もし両群で同様の結果が得られたとしたら、アスベストと喫煙の肺癌発生機序が違うこととは関係なしに、ベータカロチンとレチノールは肺癌発生に影響するといえます。他のハイリスク群や健康人にも結果をあてはめることができるかもしれませんが、もし2群の間で差があれば、発癌機序と予防に関してより掘り下げて研究するべきです。つまり例えば喫煙とベータカロチンの間に発癌発生機序の間でリンクする部分があるかもしれないと考えられます。一方、2群間の差は特別の集団にあてはまるベータカロチンの効果と考えられますから、他のリスク群や健康人に対しては結果をあてはめにくくなります。更に現実的なり点として、いくつかの暴露群を選択する理由はより多くの試験参加者を募る事ができるからです。

5. 治療群とプラシーボ群の間で肺癌の診断に違いがあった可能性はありますか？ベータカロチンとレチノールをよく使用する population で調査できると仮定して、純粹な観察研究と介入研究とではどのような違いがありますか？

治療法に関しては blind で行なっています。極一部の参加者に手が黄色くなるといった兆候が認められていますが3%であり、無視できます。よって与えられた服薬は治療群とプラシーボ群で同程度守られていたと考えられます。診断の確認に関しても治療群とプラシーボ群の間で差があったとは考えられません。しかし指が黄色くならないまでも、尿に色がついたりすると参加者は自分はベータカロチンを飲んでしていると自覚したのではないのでしょうか？さらにそのようなことを診察する医師にそっと耳打ちしているか

もしれません。

一方観察的研究では、ビタミン剤を自発的にとる人々は軽微な症状で医療機関を受診し、肺癌発見も早いかもしれません。あるいは逆に、ビタミン剤をとることにより、予防効果を過信して医療機関受診から足が遠のくかもしれません。いずれにしても、confounder や bias を避けることはできません。

本研究の場合、結果に影響する程、blind level は破られていないと想像されますが、治療によって患者さん(あるいは参加者)が本当の治療であるかプラシーボであるか気が付いてしまう場合blinding は失われます(treatment-related loss of blinding)。AIDS における AZT 臨床試験で、用量を増やしたとき、骨髄抑制をきたすため輸血が必要となります。よって患者さんの一部は自分が治療薬に振り分けられたことを知ってしまいます。このことにより患者さんが服薬を避けたとしたら、最終的 AZT の効果(副作用)が過小に評価されることとなります。

6. どの期間、肺癌の発生頻度は治療群よりプラシーボ群で高かったですか？

18 ヶ月以降で治療群がプラシーボより肺癌発生が高くなっています。しかしよくみると4-6年にかけて2つの群は近付く傾向をみせています。治療効果が6年で判定できるかどうかは疑問であり、もう少し長期に経過観察してもよかったのかもしれません。難しいところです。

7. 途中解析によって介入研究を中止することがあります。観察研究でも同様のことがありえますか？

観察的研究であっても、予定より研究を早期に中断することもありえます。仮に健康に関するニュースが study population の行動を変えたとしたら、研究を継続しても意味がなくなるため中止せざるを得ません。科学的ニュースが伝わっても、study population の生活が変わらなければ研究は継続できます。また工場での疫学調査など、観察している暴露因子が工場の方針などによりコントロールされる場合、randomization と似た状況になり、中止する場合もありえます。

アメリカで生活すると日本の食生活がいかにバランスよいか実感します。腹八分目とか、多品目のおかずを食べなさい、とかお肉をたべたら野菜も多く、よくかんで、食後はしばらくじっと座っていなさい、食べて直寝ると牛になるよ、といったことを子供のころから教えられたものです。アメリカ人はコレステロールが悪いとなると、乳製品を含むほとんどの食品から脂肪を除去してしまう。人参をたくさん食べると肺癌になりにくいかもしれないと聞くと、ベータカロチンが肺癌を予防するかもしれないと直感してしまう。最近ブロッコリーを食べると大腸癌になりにくいとか食事内容と成人病の問題が注目を集めています。科学的根拠はありませんが、日本の昔からあるおしえは以外と核心をついているのではないのでしょうか？

以下の論文もベータカロチンの癌予防効果をランダム化臨床試験にて検討しています。対象は健康な女性ですが、やはり他の結果をみて2年で中止しています。以下は筆頭著者である Lee 先生の講演より引用しました。

「ベータカロチンと癌・心血管疾患：女性健康研究より」

$\beta$ -Carotene Supplementation and Incidence of Cancer and Cardiovascular Disease: the Women's Health Study.

Lee I-M, et al. J Natl Cancer Inst 91:2102- 2106.

## 要旨

(背景) 観察的研究ではベータカロチンを含んだ野菜や果物を多く摂ると癌の発生が少ないと言われています。しかしランダム化臨床試験ではベータカロチンを投与しても、予防効果は無いが、むしろ悪いという結果が報告されています。本臨床試験は女性に対するベータカロチンの効果を検討する目的で行なわれました。(方法) The Women's Health Study はランダム化、二重盲検、プラシーボコントロールを用い、癌と心血管疾患に対して予防効果があるか否かについて 45 歳以上の女性 39,876 人を対象にアスピリン、ビタミン E、ベータカロチンを投与してそれぞれの効果について検討しました。しかしベータカロチンに関しては、2.1 年で中止となりました。統計テストは two-sided で行なわれました。(結果) ランダム化にてベータカロチンを服用した女性 19939 人とプラシーボ 19937 人を比較したところ癌の発生(incidence)、心血管疾患、全死亡率に関して差を認めませんでした。ベータカロチン群では 378 人の癌患者が、プラシーボ群では 369 人の癌患者がありました。

## 方法

### デザイン

Women's Health Study はそもそもアスピリン、ビタミン E、とベータカロチンの癌と心血管疾患に対する予防効果(primary prevention)があるかどうか検討するためにはじまった、randomized, double-blind, placebo-controlled trial です。そもそもの構想は 1986 年頃からあり、グラントの取得等もあり 4 年後の 1990 年より研究をスタートしています。本研究の開始した理由は、既に physicians study あるいは doctors study が先行していましたが、女性はホルモンをはじめ諸々の点で男性と生物学的に異なる側面をもっており、男性にあてはまる事実は必ずしも女性であてはまらない可能性もあつ



たからです。また非常に多くの人々から協力を得るためには、研究も目的をより深く理解し、最後まで多くの人に服用ルールを守ってもらうためにヘルスクエアに關与する女性を選択しました。その理由は generalizability (external validity)を落とすことになりませんが、internal validity をあげることの方が本研究においては重要であると考えられたからです。まずは 1,757,247 人に質問用紙が送られました。そのうち 453,787 人から質問用紙が返却され、この臨床研究に参加する意思があり、かつ癌・心疾患・脳血管疾患を持たない女性は 65,169 人でした。1993 年 4 月より以下の 8 種類の治療コースにランダムに振り分けました。3 種類のカプセルのうち、2 つは薬、1 つはプラシーボ、1 つは薬 2 つはプラシーボ、3 つともプラシーボの組み合わせをつくりました。100mg のアスピリン、600IU のビタミン E、50mg ベータカロチンを交互に投与します。ベータカロチンに関しては 19939 人はベータカロチン、19937 人はプラシーボにランダムに振り分けてあります。喫煙者はそれぞれ 13%ずつで同じでした。最終的に参加してくれた女性の数がいかに減少してしまっただかがわかります。さらに全ての参加女性からは書面にて同意書 (informed consent) をとって randomize した女性は 39,876 人でした。臨床試験計画は Brigham and Women's Hospital における Institutional Review Board にて承認を受けた後、その後 US Department of Health and Human Services から承認を受けています。また Minority Recruitment も重要な問題として取り上げ、minority nursing boards とコンタクトをとったり、Puerito Rico も含め、minority が多く読む雑誌に宣伝するなどしました。

National Cancer Institute および Data and Safety Monitoring Board of the Women's Health Study がそれぞれ独自に討論し、1996 年 1 月 18 日ベータカロチンに関する臨床試験を中止するに至りました。そもそも data and safety monitoring board は参加者の安全を確保するために組織されるわけですが、試験が始まる前に stopping role を創り、臨床試験自体には参加しません。定期的にその臨床試験を検討し、途中経過をみながら試験の継続ないし中止を示唆することができます。Physicians' Health Study において 22071 を対象としたランダム化臨床試験で「ベータカロチンの使用は発癌発症に影響しない」という結果が先に判ったのが主な理由です。また他の 2 つの臨床試験ではベータカロチンはむしろ悪影響であるという結果が発表されました。そこでベータカロチンのみ中止し、他のアスピリンとビタミン E に関しては継続しています。

## 治療と経過観察

初年度は6ヶ月毎、それ以降は1年に1回質問用紙を送り、治療をどの程度守っているか、end point が発生していないか、副作用が発生しないかを尋ねています。また結婚により名前が変わったりすることもあるので、他にコンタクトを取れる人をあげてもらおうようにしました。ベータカロチン治療は平均2.1年(0.0 - 2.72年)で中止となりましたが、ベータカロチン中止の時点では、87%の治療群が少なくとも十分量を服用していることが質問用紙の回答より判っています。現実的に、およそ年間5%の割合で compliance が失われるようです。一方プラシーボ群の9.9%は市販のビタミンAあるいはベータカロチンを含む薬を飲んでいました。ベータカロチンの投与を中止した後、2年間経過観察し、即ち投与開始から平均4.1年後に諸々の end points を測定しています。病気等の情報は99%の参加者よりえることができました。

自己申告がどの程度正しいかを判定するために1995年9月から10月にかけてボストン界隈在住48人の女性のベータカロチンの血中濃度を測定しています。測定日の3日前に参加者に電話連絡し、血中濃度測定に関しては知らせませんでした。私はベータカロチンの半減期を知りませんが、電話を受けてから3日間多めにベータカロチンを摂ったらやはりプラシーボより血中濃度が上がるのではないかという気もいたします。結局ベータカロチンの血中濃度は0.70 mg/dl でプラシーボ群の0.20 mg/dl より有意に( $p < 0.001$ )高く、Spearman correlation coefficient においても0.86 ( $p=0.04$ )でした。特に長期の臨床試験において、その因子に対する知識や認識が変わったり、参加者の初心が徐々に失われていくため、時間とともにバイアスを生じます。

## End points の確認

死亡に関し得は家族からの申し出、郵便局からの返信により知りました。そして遺族に書面をもってカルテをみる許可を申し出ました。カルテは病院または主治医より得ました。癌、心血管系、死亡に関する報告は、治療法を知らない(blind) end point committee of physicians により確認されました。カルテ等の客観的情報を得ることができなかった場合には、end point を認めなかったものとして扱いました。最終的に72%の end point に関してカルテで確認することができました。97%の癌については病理所見により確認しています。残りは、腫瘍マーカー、レントゲン等を用いた臨床データに基いています。非致死的心筋梗塞に関してはWHOの基準を用いました。

## 統計学的解析

randomization したものは全て検討に含めるのを原則とし、randomization によるグループ間のバランス、Cox proportional hazards regression を用いて時間要素を含めた結果を解析しました。

## コメント

最初の論文はベータカロチンは肺癌発生を上げると結論し、2 つ目の論文は癌発生、心血管疾患、全死亡に関して有意差なしとしています。最初の女性ホルモン製剤と心筋梗塞の関係において、健康な人を対象とした場合と既に心筋梗塞の既往のある人を対象とした場合とでは異なる結果を得ました（もちろん cohort vs. randomized study の相違もあります）。ここでも、最初の研究はアスベスト暴露者と喫煙者という肺癌になりやすい人々を選んでいきます。また男性の方が多くなっています。一方後半の論文は健康な女性を対象としています。また調査方法も異なります。つまりこの2つの論文の異なった結論に対して、一概にどちらがあっていて他方が間違っているということとはできません。非常に多くの対象を用いた臨床研究でも、一般論に適用したり、自分の目の前にいる患者さんに適用する際には慎重に論文を読まなくてはならないといえます。現在日本では EBM という言葉が流行っていますが、中途半端に論文を理解するとかえって患者さんに逆の治療をしかねない危なさがある点を十分知っていただきたいと思います。