

## 特集

婦人科がん臨床試験参加に必要な知識

## 2. がん臨床試験の生物統計学

山本精一郎\*

国立がんセンターがん対策情報センター/JCOG データセンター。

## 要旨

Phase1 臨床試験の目的は、試験治療の最適用量を決めるために毒性によるスクリーニングを行うことである。Phase2 試験の目的である Phase3 へ進むかどうかの決定には、毒性の情報やほかの開発薬剤の状況なども加味するので、統計的判断は有効性の足切りの目安として使うのが一般的である。Phase3 試験の論文を読む際には、primary endpoint に関する結果が重要であり、有意かどうか、どこまでいえるか（何が検証的で何が探索的か）を理解しておくことが必要である。

Key Words 生物統計学, 臨床試験デザイン

## はじめに

本稿では、がんの臨床医ががんの臨床試験の論文を読むのに必要な知識、おもに統計学的側面について知っておくべきことを述べる。これらはもちろん、臨床試験に参加する場合にも知っておくべき知識であり、自ら臨床試験を計画する研究者はもっと深く知る必要がある。また、論文を読むには統計だけでは不十分なので研究デザインについて理解することが必要であるため、デザインについても説明する。逆に、統計の概念自体を本稿のみで説明するのは難しいため、参考図書も合わせて学習していただきたい。

## Phase1 臨床試験論文を読むのに必要な統計

Phase1 臨床試験の目的は、試験治療の最適用量を決めることである。細胞障害性抗がん剤の場合には、毒性と有効性が比例することを利用して、毒性によるスクリーニングを行うことによって、推奨用量を決定してきた(図1)。この目的のために最もよく用いられるデザインは3例コホートである。これは、3例ずつ同一用量を投与し、毒性で評価した最大耐用量となるまで用量を増加する方法である(図2)。3例コホートデザインは、それを用いて実際に臨床試験を計画し、実施する際にも、また、3例コホートを用いた臨床試験論文を読むためにも統計的な知識は特に必要ない。3例コホートは、2/6の対象者に用量制限毒性(Dose Limiting Toxic-

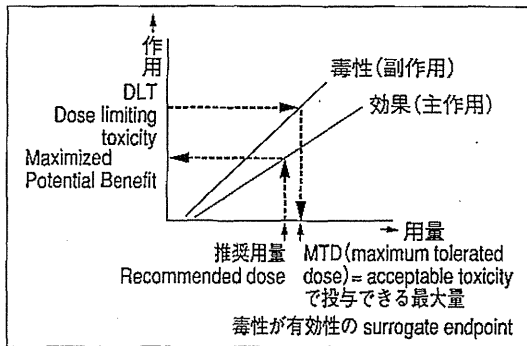


図1 細胞障害性抗がん剤 p1 デザインの rationale  
“More is better” with acceptable toxicity

ity : DLT) が発現すると中止するようなデザインとなっているため、抗がん剤としては DLT が 33% あたりが限界ラインと想定しているのかもしれないが、統計的に考察すると、実際には 33% よりも早めに中止となることが多いと考えられる。3 例コホートに統計的な根拠はあまりないが、この方法でこれまで開発を行ってきて、特に問題が生じず、やりやすくわかりやすいからということ、現在でもよく用いられる経験則的な方法であるといえる。3 例コホートは細胞障害性抗がん剤のために開発されてきたデザインであるが、分子標的薬のように、必ずしも毒性と有効性が比例するとは限らない場合にも、(明らかに優れた方法がほかにないため) 標準的に用いられているようである。

3 例コホートはかなり低い用量から開始し、順に増量するため、なかなか最適用量に達しない、最適用量付近でのサンプルサイズが小さい、といった問題が指摘されている。そこで、この点を解消するために提案されている方法が CRM (continuous reassessment method) である。この方法は、1 例ごとに統計的に用量反応曲線を推定し、それに基づいて最適と思われる用量を投与、その結果に基づいて用量反応曲線を推定し直し、次の対象者の用量を定める、を繰り返す方法である。ベイズ流統計学の手法を用いているが、3 例コホートに比べ最適用量付近が

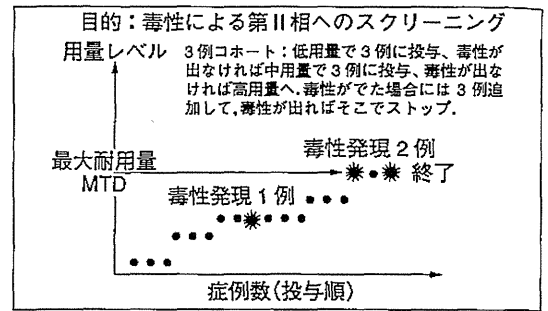


図2 第I相臨床試験の典型的なデザイン  
[By courtesy of Dr. Ishizuka]

選択されやすい優れた方法であることを知っていればよい。実際に実施するには、3 例コホートに比べ、手順がかなり複雑になることもあって、それほど頻繁には用いられていない。

## Phase2 臨床試験論文を読むのに必要な統計

Phase2 臨床試験の目的は、試験治療の開発を続け、標準治療を対照とした第 III 相試験 Phase3 を行うかどうかの判断を行うことである。Phase3 へ進むかどうかの決定には、毒性の情報やほかの開発薬剤の状況なども加味するので、統計的判断は有効性の足切りの目安として使うのが一般的である。シングルアーム試験やランダム化選択デザインなど通常 100 例以下のデザインが用いられることが多いが、検証的な結果ではなく、いずれにしろ、その先に phase3 試験が必要であることを理解すべきである。

おもなデザインは、閾値・期待値を用いたシングルアームのデザインである。これは、primary endpoint (奏効割合 response rate など) に対して、それ以下なら開発を中止すると考える値を閾値として設定し、真の response rate がある値 (期待値) であった場合には、実際に得られるデータによって閾値以下であることが高い確率で統計的に棄却できるようにサンプルサイ

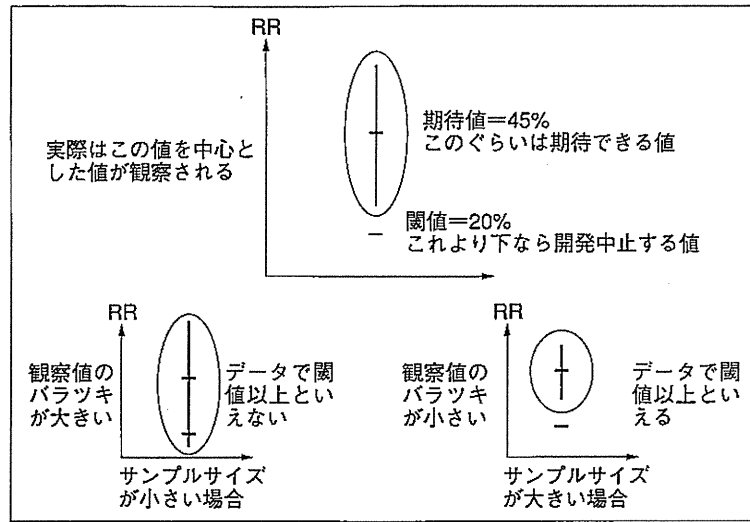


図3 閾値、期待値を用いたサンプルサイズ設計

ズと decision rule を設定するデザインである (図3)。閾値以下であるということが棄却されれば統計的には有効性が期待できるということになる。このデザインのオプションとして、試験の最後まで待たずに、途中ですでに無効が証明された場合には登録を終了する 2 stage design が用いられることも多い。

phase2 臨床試験の論文を読む場合に必要なのは、試験に実際にエントリーされた対象者の成績が、対照としていた治療の成績と比較可能かどうかをチェックすることである。たとえば、ヒストリカルコントロールとの比較の場合、実際に試験にエントリーされた対象者とヒストリカルコントロールの対象者がどのくらい比較可能かをチェックすることが重要である。適格規準に記載された対象者が必ずしも満遍なく試験にエントリーされるわけではないため、実際に試験に登録された対象者がどのような属性の分布をもっているか、それがヒストリカルなデータとの間で大きく異なっていないかを検討する必要がある。特に time-to-event (生存や無増悪生存期間) をエンドポイントとする場合には、ヒストリカルコントロールのデータがよほど

stable でないと比較できない。time-to-event のデータは、試験に組み入れられた対象者がどういふ集団かによって大きく異なるからである。より比較可能性を担保するために、ヒストリカルコントロールでなく、Phase2 試験であっても、ランダム化した対照群を設定する場合がある。この場合に注意すべきなのは、ランダム化した対照群があったとしても、サンプルサイズが小さいので検証的でなく、あくまで Phase3 へのスクリーニングと考えることである。

対照群を置くためではなく、新治療同士をランダム化比較するランダム化選択デザイン (randomized selection design) が用いられることもある。これは、患者を2つ以上の試験治療にランダムに割り付け、最も高い奏効割合が得られた治療を選択し、Phase3 試験の候補とするものである。つまり、2つ以上の試験治療がある時に優先順位をつけるデザインといえる。このデザインでは、非常に promising な治療法があった場合、その治療法を高い確率で選択できるようにサンプルサイズと decision rule を決定するものである。図4の例のように、真の response が Drug A で 35%、Drug B で 20%

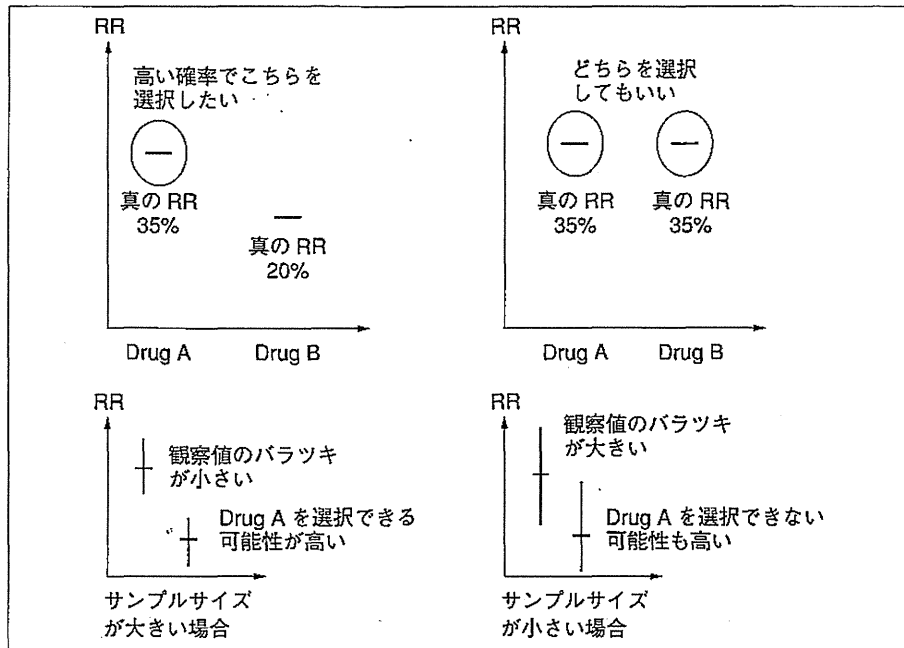


図4 randomized Phase2 デザインの考え方

である場合には、Drug A を 90% の確率で選択できるためには 29 例が必要となる。このデザインで注意すべきことは、どちらも同じ有効性をもつ治療法の場合には、50% の確率で強制的にどちらかを選択してしまうことである。差がない場合にどちらかに優劣をつけることが間違いとするならば、 $\alpha$  エラーが 100% といってもいいだろう。すなわち、まったく検証的試験とはいえないデザインであり、有効性に大きな差がない場合には、有効性の高い治療法を正しく選べない確率も高くなってしまふ。

繰り返すが、Phase2 臨床試験の結果によって、Phase3 へ進むかどうか (閾値・期待値)、どのレジメンを選択するか (選択デザイン) の決定するには、毒性の情報や他の開発薬剤の状況なども考慮する必要があり、統計的判断は有効性の足切りの目安として使うことが一般的である。サンプルサイズの目安として、シングルアームだとだいたい数 10 名 (<50) であることが多く、患者数が少ない疾患などで、次に

phase3 を実施することが現実的に不可能で、やむを得ず Phase2 試験に検証的な意味をもたせたい場合には 100 名以上とすることもある。ランダム化比較している場合には、だいたい 100 名程度であり、ランダム化選択デザインだと 100 名以下であろう。いずれにしろ、検証的な結果ではないことを理解することが大切である。

### Phase3 臨床試験論文を読むのに必要な統計

Phase3 の論文を読む際には、まず abstract で研究の概略と質を把握することが効率的である。研究仮説は何か、primary endpoint は何か、デザインはどうか、試験が予定通り完遂しているか、などが評価のポイントである。研究仮説として、比較する治療は何と何か、どちらの治療が標準としているかを把握する。適切でない治療を標準治療としてしまっている試験は結果の解釈が困難である。次に、primary endpoint として最も確実である全生存期間 (overall sur-

vival: OS), すなわちあらゆる死因による死亡までの期間, を用いているか(死因別死亡はあまり使われない)を調べる. progression-free survival (PFS), disease-free survival (DFS) や time-to-treatment failure (TTF) を用いている場合はそれらがその疾患領域で確立されているエンドポイントであるか知ることが重要である. そうでなければ, 必ずしも検証的な研究結果としてとらえてはいけないことになる. デザインは優越性試験か非劣性試験か, 新しい治療が toxic new なら優越性, less toxic new なら非劣性デザインを使う必要がある. 非劣性試験の場合, 臨床的に意味のある secondary endpoint が定義されていて, それが標準治療に勝っていることが必要となる. 試験が完遂しているかも大きなポイントである. 予定通りのサンプルサイズで終了しているか, 途中終了の場合, 理由は何か, また, 追跡は十分かなどもチェックするポイントである. abstract にこれらの記載がない場合は, 本文の中でチェックすることになる.

introduction では, 研究を行うに至った背景, 特に対象と標準治療の決め方, なぜその新治療が有効であると考えられるかの説明の妥当性を見る. それらを踏まえ, どのようなデザインが適当かを自分で考えてみることはよい勉強になるであろう.

methods では, どのような試験グループが行っているか, eligibility の記載は正確か, 統計記述は十分かなどをチェックする. たとえば, sample size 設計の根拠がきちんと書かれているかといったところである. 優越性試験であれば, 期待する治療効果が大きすぎないかをチェックすべきであるが, これは研究の結果, 統計的な差が認められなかった場合に検出力不足のせいかどうかを検討するためである. 非劣性試験であれば, 設定している許容域が広すぎるかどうかをチェックすべきで, 広すぎる場合

には, 必要以上に統計的に有意になりやすいことになる. 中間解析のやり方を含め, 解析方法がきちんと書かれているかも重要なポイントである. また, データマネジメントについて記述があるかなども研究の質を推測する手掛かりになる. 最近の論文では, CONSORT 声明に従った研究の流れ図が記載されている論文も多いが, これは途中で脱落した人がどのくらいいるか, 治療のコンタミネーションがどのくらい起きているか, Intention-to-Treat (ITT) 解析をしているか, などを容易に把握することができ, 研究の質を評価するのに有用である. 自分でこの流れ図を書いてみるのもよい勉強になる.

results では, まず背景因子の比較が示されることが一般的である. ここでは, 重要な予後因子に関して, 群間に大きなアンバランスがないかどうかを確認する. もし大きなアンバランスがあれば, 結果の解釈の際に注意すべきであるし, 統計モデルでアンバランスな要因を調整した解析結果なども参考にすることになる. 背景因子ごとに  $p$  値が示されていることも多いが, 割り付け調整因子になっていない場合には, 20に1つは統計的に有意にアンバランスになってしまうので,  $p$  値自体にあまり神経質になる必要はない. ランダム化で期待しているのは, 「同じ治療をすとした場合の予後が等しくなる」ことであって, 個々の背景因子のバランスではない. アンバランスがある場合でも, 最も重要な結果はランダム化に基づいた解析であり, 統計モデルで調整した解析はあくまで確認のための結果であると考えべきである.

背景因子の比較の次には, 有害事象・毒性や治療のコンプライアンス, 有効性の secondary endpoint の結果等が順に示されることが一般的である. もともと, 有害事象に関してサンプルサイズ設計を行っているわけではないため, 有害事象が両群で統計的に有意に異なるかどうかの  $p$  値は参考程度にすべきであって, 有意かど

うかだけで判断すべきでなく、より臨床的な知識を用いてその差を検討すべきである。治療のコンプライアンスについては、経験上、正確に定義することと、それにあったデータを収集するのが難しいため、そのことに注意して解釈する必要がある。有効性の secondary endpoint として response rate が記載されている場合も多いが、Phase3 では response rate はあくまで参考値とすべきである。なぜなら、response rate はもともと、おもに Phase2 において新治療の有効性をほかの治療法と比較するために用いることを目的としたエンドポイントであり、phase3 では全員について計測されていないことや central review されていないことも多いからである。QOL について報告されている場合もあるが、QOL 自体を primary endpoint にした研究でなければ、これも参考程度と考えるべきである。patient report の QOL が重要であることは間違いないが、測定方法やデータの収集、適切な解析など全部をクリアし、正しい結論であることを保証することが非常に難しいからである。

最も重要なのは primary endpoint の結果である。前述の通り、primary endpoint としては全生存期間 OS が最も望ましく、PFS が primary endpoint である場合でも、OS は大きな意味をもつ。用いられる統計手法は、Kaplan-Meier curve による生存曲線の比較が一般的であり、打ち切りがどのくらいあるか（ひげはついていない場合もある）がデータの maturity の目安となる。生存曲線の比較には、Logrank 検定を用いることが多い。Generalized Wilcoxon 検定を用いることもあるが、これは logrank 検定に比べ、初期のイベントに大きな重みをおいている方法である。検定結果が有意であるかどうかとともに、効果の大きさ（ハザード比や Median Survival Time (MST) の差) についても着目する必要がある。Cox 回帰は、単変量、多変量ともに、ハザード比を出すのに用いられ、前述のように、

多変量調整した結果はあくまで補助的なものと考えべきである。

discussion では、結果の解釈について議論される。優越性デザインの場合、試験がしっかり計画、実施、解析された場合には、primary endpoint で有意な結果が得られた場合、新治療が勝っていると判断する。その際、効果の大きさについても考慮することが重要である。primary endpoint で有意な結果が得られなかった場合には、対象とした集団全体では新治療が勝っているとはいえない。この場合にも、探索的にサブグループ解析をして、次につながる仮説を立てることが重要である。また、論文を読んでみて、試験がしっかり計画、実施、解析されていない場合と考えられる場合には、結果の解釈に注意が必要である。サンプルサイズが小さい、途中でリクルートできず終わった試験などは、有意であっても偶然の可能性も高いため、探索的な試験として考えるほうがよいといえる。

非劣性試験の場合には、結果の解釈に注意が必要である。非劣性試験とは、試験治療の primary endpoint (たとえば OS) が、標準治療と比べて「許容できる差」以内にあるかどうかを検証するデザインである(図5)。許容できる差以上に劣っている、という仮説を棄却することによって、許容できる差以内であることを検証する。この場合、有害事象が少ないなど、ほかにメリットがあるからこそ、OS は同等であればよいといえる。実際には、OS で同等以上が期待できる新治療に対し、ほかにメリットがあるために非劣性デザインを組むことが多い。非劣性デザインでは、試験がしっかり計画、実施、解析された場合には、primary endpoint で有意な差、かつ secondary endpoint で新治療の有効性が示されている場合、新治療が優れている、あるいは標準治療のオプションになりうると判断することになる。この場合、非劣性の許容域が

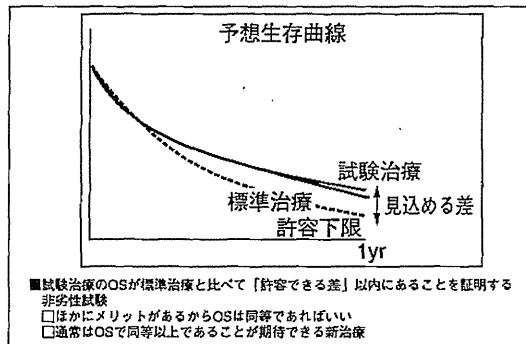


図5 非劣性試験とは

広すぎないこと（5年生存率で10%やハザード比1.5など）に注意する必要がある。primary endpointでは有意であったが、secondary endpointで新治療の有効性が示されていない場合、標準治療のオプションになり得るかは判断が難しい。また、primary endpointで有意な結果が得られなかった場合には、新治療は新しい標準治療、もしくはそのオプションとはなれないと考えるべきである。ただし、信頼区間で解釈して許容域を再定義することは解釈のうえで有用である。たとえば、5年生存率で3%劣っていないという許容域は棄却できなかったが、5%は劣っていないことは信頼区間による検討からいえる場合がある。この場合、5%に臨床的意義がある場合には、ある程度有効であるという判断をしてもよいであろう。もし、試験がしっかり計画、実施、解析されていないと考えられるような場合には注意が必要である。なぜなら、非劣性試験は、きっちり行わないと有意になりやすい（非劣性が証明されやすい）からである。これは、治療の不遵守で両群が同じ治療をしたり、測定の変異があると差が消えていって同じ成績に近づいていくからである。

試験のprimaryな結果に加えてdiscussionでよく議論されるのは、サブグループ解析の結果である。サブグループ解析をしているいろいろな検討をすることは非常に重要なことであるが、結

果の解釈には慎重になるべきである。2007年に出されたNew England Journal of Medicineのサブグループ解析の報告に関するガイドラインが解釈の参考になる。これには、サブグループの結果をabstractに報告してよいのは、それらがprimary endpointに対するものである場合、かつ、事前に規定したサブグループ解析を解析全体として解釈できる場合に限る、としている。また、methodsにサブグループ解析の数（報告数、実施数）を明記すること、事前に規定して実施したものとそうではなく事後的に実施したものを区別すること、エンドポイントと解析法の明記、特に興味のあるものをhomogeneityの評価と区別して記載、 $\alpha$ エラーへの影響の程度とそれに対する対応法の明記、治療効果の不均一性の評価には交互作用に対する検定を行うことが規定されている。discussionでは、過度の解釈を避ける、信憑性の評価、限界を記載、結果を支持/矛盾する先行研究結果を提示するなど、細かい規定が設けられている。一読することを勧めたい。

Phase3論文を読む際の注意をまとめると、abstractでは、研究の概略と質（研究仮説は何か、primary endpointは何か、デザインは適切か、試験が予定通り完遂しているかなど）を把握し、introductionでは、研究を行うに至った背景、特に対象と標準治療の決め方、なぜその新治療が





## ● 疫学・予防

## 乳癌疫学の最近の動向

\* 国立がん研究センターがん対策情報センター \*\* 同 室長

溝田 友里\* 山本 精一郎\*\*

## 要旨

我が国における乳癌の動向は、女性乳癌死亡者数が 11,797 人、年齢調整死亡率 11.9 人 (2008 年) であり、女性乳癌罹患数の推計は 50,549 人、年齢調整罹患率は 62.0 人 (2004 年) である。女性のがんでは、年齢調整死亡率が 2 位、年齢調整罹患率では 1 位であり、いまだ一貫した増加傾向にある。乳癌リスクファクターとして、乳癌家族歴や家族性乳癌遺伝子 1, 2 (BRCA1, BRCA2) などの遺伝的素因、初経や閉経に関する内分泌環境因子、出産や授乳など社会環境と共に、予防が可能な生活環境要因として、肥満や身体活動、アルコール摂取などが明らかになっている。乳癌の増加に早急に対応するために、まずは利用可能なエビデンスの中から、予防行動につなげていくことが重要である。

## 乳癌の罹患率と死亡率

乳癌は世界的にみても女性の最も多いがんであり、国際がん研究機関 (International Agency for Research on Cancer : IARC) の推計によると、2002 年に新たに 115 万人が罹患している。また、女性のがんによる死亡の第 1 位となっており、2002 年における女性乳癌死亡者数は 41 万人であった<sup>1)</sup>。

がん対策の立案と評価には、その基礎となるデータとして地域、あ

---

キーワード：罹患率，死亡率，リスクファクター，疫学，予防

るいは国レベルでのがんの罹患と死亡の把握が不可欠である。そのため、我が国における乳癌罹患率、死亡率について、どのようなリソースが存在し、どのような傾向があるか知ることは乳癌の対策を行ううえで重要である。

### 1. 死亡率

我が国では、がんの死亡動向は厚生労働省の人口動態調査によって全数把握されている。人口動態調査は明治時代から実施されている政府統計であり、国際的にみても精度が高く、また公表時期も調査年から1年遅れと早い。人口動態統計によるがん死亡データ（1958～2008年）ならびにそれを用いた種々のグラフは、国立がん研究センターがん対策情報センターのがん情報サービス（<http://ganjoho.jp/professional/statistics/index.html>）より入手可能である。

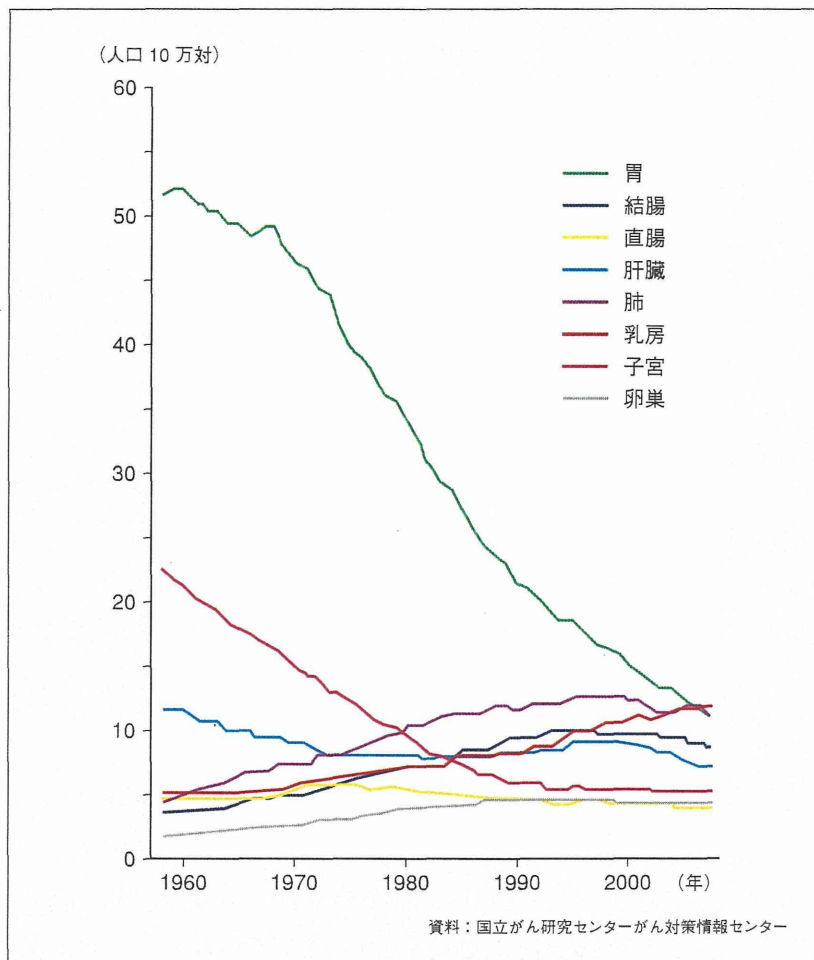
がん年齢調整死亡率の推移をみると、全がんでは男女共1990年代後半から減少傾向にあり、部位別では、男女共に胃癌、結腸癌、直腸癌、肝臓癌などで近年減少傾向がみられる。しかし、女性の乳癌については、1960年代以降一貫した増加傾向にあり、近年でも明らかな増加がみられる（図1）。2008年の女性の乳癌死亡者数は11,797人であり、年齢調整死亡率は、大腸に次いで高く、人口10万対11.9人である。年齢別死亡率は50代まで直線的に増加し、その後は80代まではやや減少する（図2）。また、女性乳癌死亡率を出生年代別にみると、最近の罹患率の上昇（後述）を反映し、最近の出生者ほど死亡率が高くなっている（図3）。

欧米諸国の乳癌年齢調整死亡率は、日本に比べかなり高い。しかしながら、1990年あたりをピークに減少傾向が始まっており、日本との差は縮まる傾向にある<sup>2)</sup>。

### 2. 罹患率

我が国では、死亡と異なり、がん罹患に関するデータを国として系統的に把握する仕組みがない。罹患率を推定するためには、ある集団を設定し、その集団で一定期間に発生した罹患数を把握する必要があり、これを実現するためには地域がん登録が不可欠である。地域がん登録は、がん対策の立案と評価や、がん医療の向上のために世界の多くの国や地域で行われており、日本では、1950年代に宮城県、広島

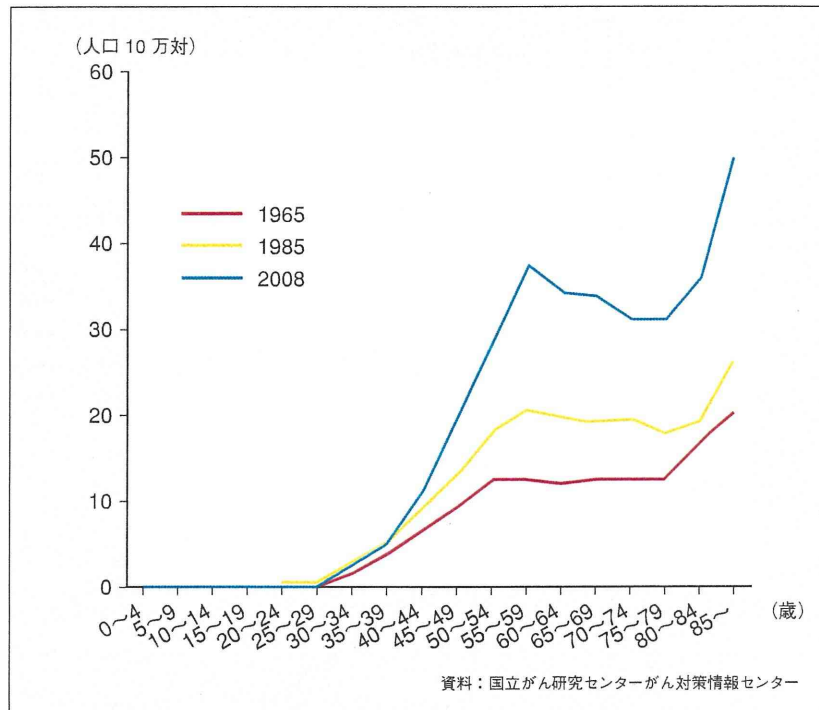
図1 部位別がん年齢調整死亡率の推移 (主要部位) [女 1985年～2008年]



市，長崎市で開始された。次いで 1960 年代に大阪府，愛知県などで始められ，2010 年 3 月時点で，日本では 36 道府県および 1 市で地域がん登録事業が実施されている。

我が国では，全国規模での地域がん登録が行われていないため，一定の精度基準を満たした 10 数府県の地域がん登録のデータに基づき，全国推計値を算出することで国レベルのがんの罹患状況を把握している。1975～1994 年の全国がん罹患推定は，厚生労働省がん研究助成金による「地域がん登録精度向上と活用に関する研究」班が，1995 年以降の推計は，厚生労働科学研究費補助金第 3 次対がん総合戦略研究事業「がん罹患・死亡動向の実態把握の研究」班で罹患データを収集

図2 年齢階級別死亡率 複数年（女性）[乳房 複数年]

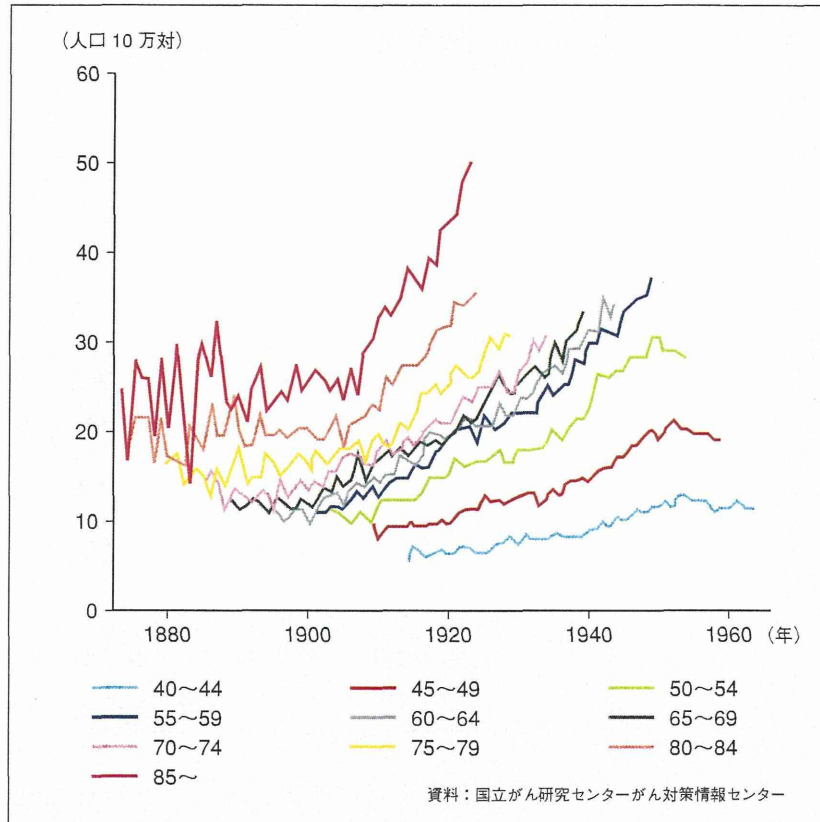


し、国立がん研究センターがん対策情報センターで解析・公表を行っている。集計結果は、前述のがん情報サービスホームページの集計データのダウンロードで公開されている (<http://ganjoho.jp/professional/statistics/index.html>)。各地域からのデータ提出と集計作業に時間がかかるため、公表時期は罹患年より5～6年遅れとなっている。2010年4月現在、全国がん罹患推計の最新年は2004年である。

がん年齢調整罹患率の推移をみると、全がんでは男女共1990年代前半までは増加し、その後は横ばいで、2000年前後から再び増加傾向にある。男性では胃癌、肝臓癌などで近年減少傾向がみられるが、直腸癌、前立腺癌などは増加傾向にある。女性では、胃癌、肝臓癌などで近年減少傾向がみられるが、直腸癌、肺癌、乳癌、卵巣癌で増加傾向にあり、特に乳癌と卵巣癌は1975年から一貫した増加傾向が続いている(図4)。

女性のがんの年齢調整罹患率は乳癌が最も高く、人口10万対62.0人であり、2004年の推計によると50,549人が罹患している。年齢別

図3 年齢階級別死亡率 出生年（女性・40歳以上）[乳房]

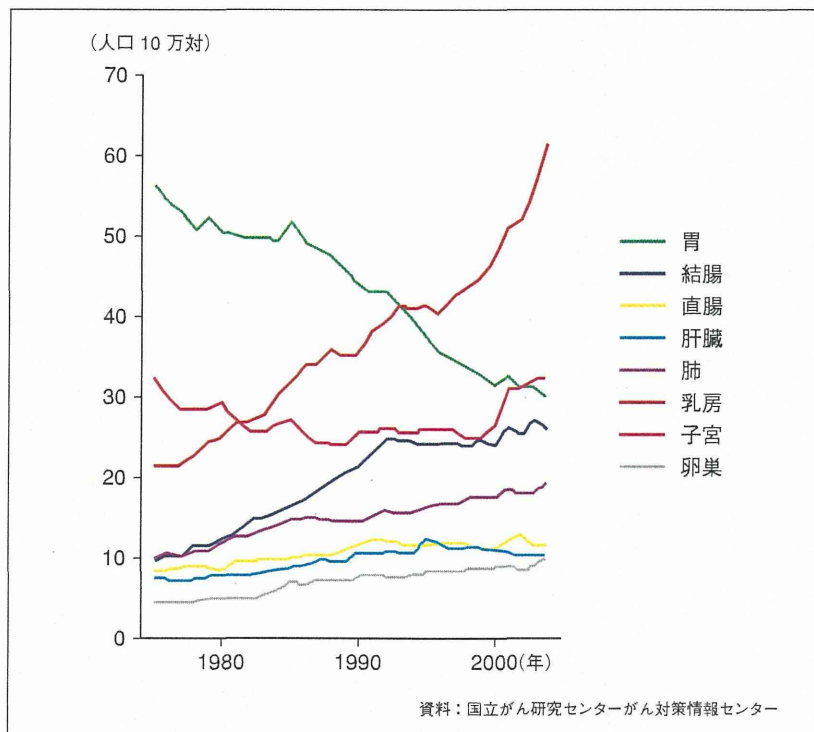


横軸に出生年、縦軸に各年代ごとの死亡率をとり、同じ年代の死亡率をつないだもの。出生年（横軸）が最近になるほど、死亡率（縦軸）が高くなっている。

にみた女性の乳癌の罹患率は30歳代から増加し始め、40歳代にピークを迎え、その後は次第に減少する（図5）。出生年代別にみると、死亡率と同様に、最近の出生者ほど、罹患率が高くなっている（図6）。

世界的な年齢調整罹患率の年次推移の傾向をみると、東アジアの人々の乳癌罹患率は、ヨーロッパ人や米国白人に比べ、一貫して低い。日本と同様、中国などの東アジア諸国でも明らかな上昇傾向を示している。ヨーロッパ諸国は同じレベルの年齢調整罹患率であり、上昇傾向を示しているが、最近一部の地域では減少傾向もみられている。米国でも、ほとんどの人種で継続的な上昇傾向にある。米国では白人が最も高い罹患率を示し、次に黒人であるが、日系人の罹患率も急速に増加し、日本に住む日本人よりも高い罹患率を示している<sup>3)</sup>。米国

図4 部位別がん年齢調整罹患率の推移（主要部位）〔女 1975年～2004年〕

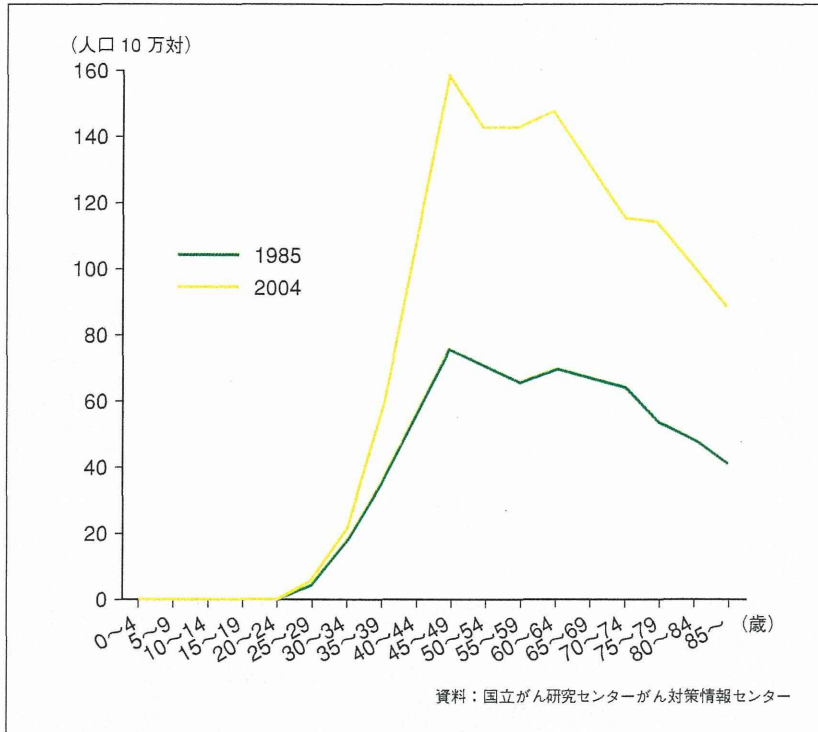


のがん罹患データは SEER (<http://seer.cancer.gov/resources/>) からダウンロードできる。また、IARC は世界各国、各地域からがん罹患情報を収集して「5大陸のがん罹患」として公表している。「5大陸のがん罹患」のデータは、世界保健機関 (WHO) がまとめている各国のがん死亡データと共に、IARC で入手が可能である (<http://www-dep.iarc.fr/>)。

### 乳癌発症のリスクファクター

前述のとおり、日本人における乳癌罹患率は 1975 年以降、大きく上昇を続けており、また、米国に移住した日系人の罹患率は、移住国の罹患率に近くなり、日本に住む日本人よりも高くなっている。これらのことから、乳癌罹患率の上昇に遺伝的要因が大きく影響しているとは考えられず、乳癌罹患率における国際的な違いは、生活習慣など環境要因の影響が大きいことが示唆される。そこで、本稿では、乳癌罹患率への影響が大きく、かつ予防可能な要因である生活習慣を中心

図5 年齢階級別罹患率 複数年 (女性) [乳房 複数年]



に、リスクファクターを紹介する。

1. 乳癌家族歴・遺伝的素因

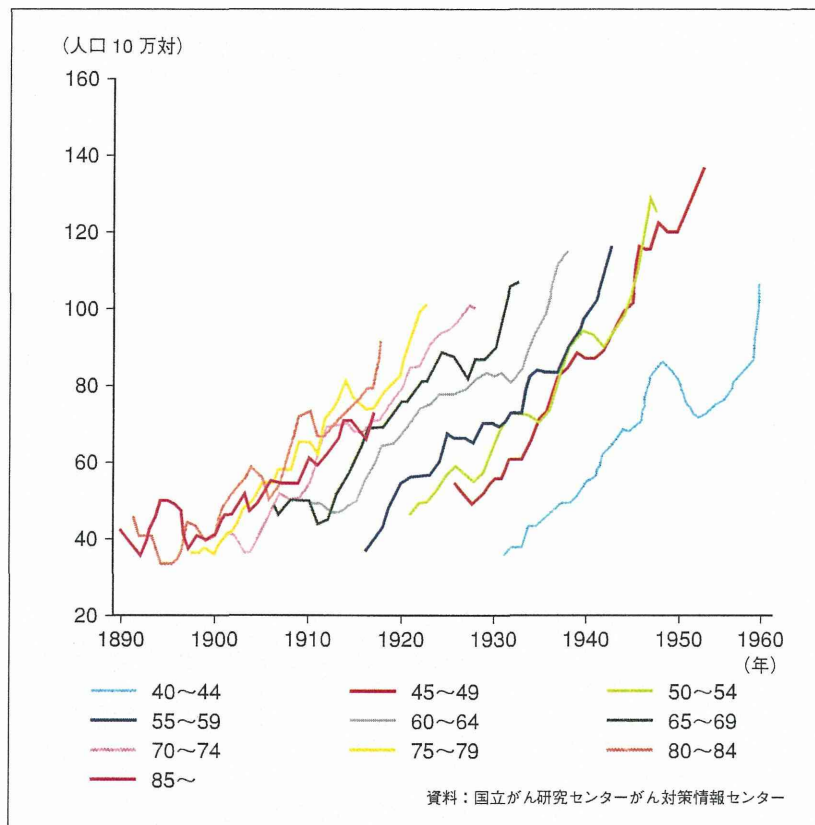
乳癌家族歴と乳癌発症との関連が多く報告されており、乳癌家族歴を有することは乳癌のリスクファクターと考えられている。

また、遺伝的素因により罹患した乳癌である家族性乳癌は、欧米のデータを中心とした推計では、全乳癌の5～10%にあたりと考えられているが、日本における正確な寄与割合は分かっていない。原因遺伝子として BRCA1 や BRCA2 が現在解明されており、これらの遺伝子変異を持つことが、乳癌リスクを上昇させる。

2. 内分泌環境因子・社会環境因子

乳癌はホルモン関連がんであり、内因性のエストロゲン暴露が乳癌リスクに関連すると考えられる。疫学的研究からも、初経年齢が早いこと、自然閉経年齢が遅いこと、出産経験がないこと、初産年齢が遅い(30歳以上)ことは、すべて乳癌で“确实 (convincing)”，あるいは“ほぼ确实 (probable)” なリスクファクターとされている。同

図6 年齢階級別罹患率 出生年（女性・40歳以上）[乳房]



横軸に出生年，縦軸に各年代ごとの罹患率をとり，同じ年代の罹患率をつないだもの。出生年（横軸）が最近になるほど罹患率（縦軸）が高くなっている。

様に授乳経験がないことも乳癌リスクを増加させることは“確実”である。

また，ホルモン補充療法は乳癌のリスクファクターである。経口避妊薬はエストロゲン単独，エストロゲンとプロゲステロンを併用した場合のいずれにおいても乳癌のリスクを高めるが，リスクは比較的小さい。これらは外因性ホルモンであり，月経や妊娠に関連する内因性ホルモンと同様に，乳癌のリスクを上昇させる。

その他，X線など，医療で用いられる電離放射線への暴露や放射線被曝も乳癌のリスクファクターである。

### 3. 生活環境因子

生活環境因子を中心とした乳癌のリスクファクターについてシステ



マテック・レビューを行い、広く世界中で活用されているものに、世界がん研究基金 (World Cancer Research Fund: WCRF)/ 米国がん研究財団 (American Institute for Cancer Research: AICR) の、食事、栄養、身体活動に関するレビューがある。その報告書である“Food, nutrition, physical activity and the prevention of cancer: a global perspective”<sup>4)</sup> は 2007 年 11 月に発表された。また、日本人に関する乳癌リスクファクターについては、厚生労働省の科学研究費補助金・第 3 次対がん総合戦略研究事業「生活習慣改善によるがん予防法の開発と評価」研究班が行っている日本人を対象にした疫学研究のレビュー<sup>5)</sup> と、日本乳癌学会が出版した『科学的根拠に基づく乳癌診療ガイドライン 5. 疫学・予防 2008 年版』<sup>6)</sup> が広く知られている。これら 3 つのレビューをまとめたものを表 1 に示す。

全般的に、WCRF/AICR に比べ、日本人を対象としたレビュー結果では、食事や栄養、身体活動と乳癌との関連を示すエビデンスレベルが低くなっていたが、全体としての傾向はほぼ同様のものだった。日本においてリスクファクターと乳癌発症との関連が弱いことの背景として、日本人を対象とした研究はいまだに少なく、乳癌との関連を判定するのに十分なエビデンスが得られていないことが理由として考えられる。また、肥満やアルコール摂取との関連が日本人では弱いことについては、日本人女性では肥満の人の割合や、飲酒習慣がある人の割合、あるいは大量に飲酒する人の割合が小さいため、対象集団内でのばらつきが小さいことが影響している可能性も考えられる。

なお、これら乳癌リスクファクターのレビューでは、男性についてのエビデンスが限られているため、女性乳癌についてのみ扱われている。

食事に関して確立されたリスクファクターは、現在のところアルコール摂取のみである。アルコール摂取が乳癌のリスクファクターであることは、閉経前後を問わずすべての年齢において“確実”である。穀類や食物繊維、野菜と果物、大豆、乳製品、脂質、糖類、ビタミン、イソフラボンなどが注目され、乳癌との関連が検討されてきたが、これらに関しては全般的に“証拠不十分 (limited-no conclusion)”である。脂肪摂取については、閉経後乳癌のリスクであることが“可能

表1 乳癌の発症と、食事、栄養、身体活動との関連（文献<sup>4-6)</sup>より引用改変）

データソース	WCRF / AICR による 閉経前乳癌	WCRF / AICR による 閉経後乳癌	厚生労働省研究班による 日本における乳癌	乳癌診療ガイドライン
成人期の身長	probable (↑)	convincing (↑)	—	—
出生時体重	probable (↑)	limited-no conclusion	—	—
体脂肪 (肥満)	probable (↓)	convincing (↑)	—	閉経前：probable (↓) 閉経後：convincing (↑)
身体活動	limited-suggestive (↓)	probable (↓)	limited-no conclusion	閉経前：limited-no conclusion 閉経後：probable (↓)
喫煙	—	—	limited-suggestive (↑)	limited-suggestive (↑)
アルコール	convincing (↑)	convincing (↑)	limited-no conclusion	probable (↑)
野菜・果物	limited-no conclusion	limited-no conclusion	limited-no conclusion	—
大豆製品	limited-no conclusion	limited-no conclusion	limited-suggestive (↓)	limited-no conclusion
ビタミンC	limited-no conclusion	limited-no conclusion	—	—
緑茶	—	—	limited-no conclusion	limited-no conclusion
脂肪摂取	limited-no conclusion	limited-suggestive (↑)	limited-no conclusion	閉経前：limited-no conclusion 閉経後：limited-suggestive (↑)

性あり”とされている。

体格については，成人期の身長の高さが乳癌のリスクになることは，WCRF / AICR の閉経後乳癌においては“確実”，閉経前乳癌においても“ほぼ確実”である。また，出生時の高体重が，閉経前乳癌のリスクであることは“ほぼ確実”とされている。

体脂肪（肥満）が閉経後乳癌のリスクであることは“確実”であり，一方で，体脂肪（肥満）が，閉経前乳癌においては予防的な効果を持つことも“ほぼ確実”である。

身体活動量が多いと乳癌リスクが低減することは，閉経後乳癌に関しては，WCRF / AICR，日本人のレビュー結果とも“ほぼ確実”であり，閉経前乳癌に関しては WCRF / AICR では“可能性あり (limited-suggestive)”である。身体活動については，月経・排卵周期や内因性ホルモンレベルへの影響を介したメカニズムも考えられているが，身体活動を行うことによって体脂肪の増加や体重の増加，閉経後の肥満が予防され，結果として乳癌のリスク低下に関連していると考えられている。

生活環境要因は，自ら行動変容できる要因であるため，予防に用いることが可能である。十分なエビデンスがあるものは多くはないが，乳癌の増加に早急に対応するために，利用可能なエビデンスの中で，ほかの疾患への影響も含め，リスクとベネフィットを考慮して予防行動につなげるべきであろう。

### 乳癌再発のリスクファクター

乳癌再発のリスクファクターについては，食事や栄養，身体活動と，乳癌の再発などとの関連を検討する疫学的研究が，日本のみならず世界的にも不足している。国立がん研究所 (National Cancer Institute : NCI) の PDQ<sup>®</sup> (Physician Data Query) (<http://www.cancer.gov/cancertopics/pdq>) では，ビール摂取や脂肪摂取，高エネルギー摂取とリスク上昇，野菜・果物摂取とリスク低下などが紹介されているが，いずれも少数の研究結果が中心で，エビデンスレベルも低い。そのため，WCRF / AICR の報告書でも明確な推奨は行われていない。

検診の普及や治療法の改善により，乳癌を抱えながら生活する人が

増えている。有効な治療法が存在しても、患者の立場からは日常生活の中で再発を防ぎ QOL を高めるために努力をしたいという思いが強い。そのようなニーズを背景として、乳癌患者における生活習慣と予後との関連を明らかにしようとする前向きコホート研究が国内外ともようやく開始され始めた。米国においては LACE Study や Pathways などが開始され、日本においては、筆者らが中心となり、厚生労働省科学研究費補助金がん臨床研究事業「生活習慣や支持療法等が乳がん患者の QOL に与える影響を調べる多目的コホート研究（研究代表者：山本精一郎）」により、乳癌患者数千人を対象とする大規模コホート研究を開始した。これらの研究の成果から、今後、乳癌の再発予防に寄与する要因が明らかになることが期待される。

#### 文 献

- 1) International Agency for Research on Cancer. Globocan 2002. (<http://www-dep.iarc.fr/>) Accessed May 13, 2010.
- 2) Kawamura T, et al: Cancer Statistics Digest. Comparison of Breast Cancer Mortality in Five Countries: France, Italy, Japan, the UK and the USA from the WHO Mortality Database (1960–2000). *Jpn J Clin Oncol* 35 (12) 758–759, 2005.
- 3) Katanoda K, et al: Cancer Statistics Digest. Comparison of Time Trends in Female Breast Cancer Incidence (1973–1997) in East Asia, Europe and USA, from Cancer Incidence in Five Continents, Vols IV–VIII. *Jpn J Clin Oncol* 37 (8) 638–639, 2007.
- 4) World Cancer Research Fund/American Institute for Cancer Research: Food, nutrition, physical activity and the prevention of cancer: a global perspective. AICR, Washington DC, 2007.
- 5) 厚生労働省科学研究費補助金・第3次対がん総合戦略研究事業「生活習慣改善によるがん予防法の開発に関する研究」[http://epi.ncc.go.jp/can\\_prev/](http://epi.ncc.go.jp/can_prev/) (Accessed May 13, 2010)
- 6) 日本乳癌学会編: 科学的根拠に基づく乳癌診療ガイドライン 5. 疫学・予防 2008 年度版. 金原出版, 東京, 2008.