



解説

がん治療研究におけるランダム化 第II相試験の意義*

山本精一郎**

Key Words : molecularly targeted agent, trial design, randomization, phase II trial

最近の薬剤開発事情

最近, 多くの薬剤が開発されており, その中には多くの分子標的薬が含まれる. これらは従来の細胞傷害性抗がん剤 (cytotoxic drug) とは異なる毒性プロファイルをもつ. たとえば, 単剤での activity が必ずしも高くない, 毒性がオーバーラップしないといった特性をもつなど, 標準的な化学療法との併用が効果的と考えられるケースがある. こういった薬剤の開発には従来の cytotoxic drug 開発とは異なる戦略が必要となってきた.

これまでの cytotoxic drug では「用量-毒性」「用量-効果」が相関関係にあり, 結果として毒性と効果が比例していたことから, 毒性による最大耐用量を推定することがそのまま効果を最大化する用量を推定することにつながるというロジックに基づいて phase I trial が行われてきた. それに対し, 分子標的薬をはじめとするいわゆる “cytostatic drug” では, かならずしも用量と効果が比例しない. そのため, 毒性の面からの最大耐用量を推定するプロセスを通して効果の面での最適用量を推定することが難しくなった. その結果, 「用量-ベネフィット/リスク比」関係を検討して至適用量 (optimal dose) を決めるという, 一般薬開発で用いられてきた考え方と同じ方法をとる必要が出てきた. つまり, 最大耐用量を探す古典的な phase I デザインでは, 至適用量が決定できなくなり, その後の開発ステージである phase II や場合によっては phase III で用量設定

を行うという必要が生じてきたということである.

また, 最近, phase III で新規薬剤と標準治療の比較を行ったが, 有効性が証明されず失敗に終わったという例がいくつかあった. 多くの phase III を実施することは通常困難であるため, 失敗した場合には, その薬剤の開発が中止されるだけでなく, 他の薬剤の開発の遅れにもつながる. 新規薬剤がそれほど多くない時代には, phase III で検討の対象となる他の選択肢がないためにこれはそれほど大きな問題ではなかったが, ほかに多くの候補薬がある現在では, phase III の失敗は治療開発全体の遅れにつながってしまう. そのため, 医療に携わるものとしてはできるだけ成功する確率の高い薬剤を選んで phase III を行いたい.

ではどうやって有効性のスクリーニングをすればよいのであろうか. 単剤で抗腫瘍効果のみられる薬剤であれば伝統的な phase II デザインで奏効割合 (response) などの腫瘍縮小効果をエンドポイントにすればよい. たとえば, CML や GIST に対する imatinib や NSCLC に対する gefitinib などがその例であろう. しかし, 分子標的薬の中には必ずしも単剤としての activity が高くなく, 標準治療との併用療法として期待されているものも多い. このような場合, phase III へ進めるべき薬剤をどのようにスクリーニングすればよいのであろうか.

このような状況下で phase II に求められる役割は, これまでの有効性によるスクリーニングとい

* Randomized phase 2 design for molecularly targeted agent.

** Seiichiro YAMAMOTO, Ph.D.: 国立がんセンターがん対策情報センター/JCOGデータセンター統計部門(〒104-0045 東京都中央区築地5-1-1); CIS, National Cancer Center, Tokyo 104-0045, JAPAN

う目的に加え、いくつかの用量や候補薬剤, combinationの中から, phase IIIで成功する確率の高い薬剤や治療法を選択するということである。この場合にrandomized phase IIが有効なデザインと考えられる。なぜなら, phase IIで標準治療のconcurrent controlと比較できれば, ヒストリカルコントロール(historical control)を用いることによる選択バイアスを防ぐことができるため, 有効性に関してsingle armの試験より高い確信が得られるからである¹⁾²⁾。ランダム化は, 新薬や分子標的薬との組み合わせで目安にすべきヒストリカルコントロールのデータがあまりない場合やresponseがエンドポイントとして適切でなく, time to progressionの方がエンドポイントとしてよい場合にはさらに有効である。なぜなら, time to progressionやsurvivalなどは, ヒストリカルコントロールと同じ選択規準で試験に組み入れられる対象者を選択しても, 各被験者がもともとつ予後によって大きく影響を受ける, すなわち, 治療以外の要因によってtime to eventに大きな違いが生じてしまう可能性がある。また, 抗腫瘍効果は腫瘍が縮小したことを直接観察できるのに対しtime to progressionやsurvivalの延長は直接観察することができない。このようなtime to eventの欠点は, 標準治療とランダム化することによって克服することができる。このように, 標準治療を含めたランダム化比較をphase IIで行うことは大きな利点をもつが, 2群を直接比較するとサンプルサイズが倍以上になってしまう。では, どのようにすればよいのであろうか。以下, Rubinsteinらの論文³⁾をもとに, randomized phase II designについてこれまでに提案されている例をあげ, その特徴を議論してみたい。

これまでに提案された randomized phase II design

1. Randomized phase II selection design

1985年にNCIのSimonら⁴⁾によって提案されたものが有名である。患者を2つ以上の試験治療にランダムに割り付け, もっとも高いresponseが得られた治療が次の研究に進むというデザインである(エンドポイントはほかのものでもよい)。2つの試験治療があるときに優先順位をつけるデザイ

ンであって, もっともよい群が高い確率で選択されることを保証するようにサンプルサイズを決定する。たとえば, 真のresponseがレジメンAで35%, レジメンBで20%のとき, 90%の確率でレジメンAを選択するためには29名が必要, といった具合である。この場合の問題点は, 「どちらも同じ場合には50%の確率でAを選択!!!」することであり, その意味で α エラーが50%ともいえるかもしれない。このデザインはdoseやscheduleを選ぶためにも用いることができる。前提として, 毒性の程度がほぼ同じことを仮定しているが, 毒性が異なる場合にもその分効果に下駄を履かせたdecision ruleを作ることも可能である。しかし, 異なる治療薬の選択の場合には本当にこれでよいのかという不安がある。というのも, 2群の治療効果がまったく同じでも, 50%の確率でいずれかの群を選んでしまうためである。また, 観察されたresponseのみで選択するため, とくに事前にどちらがよいかの根拠がまったくない場合だけ使えるともいえる。さらに, 試験の対象となる薬剤が異なる製薬企業によって開発されているものであった場合, 製薬企業がよいといえばこのデザインによる優先順位づけはできるかもしれないが, 大きなコストをかけて開発が行われていることを考えると, このような単純な判断基準が受け入れられるとは考え難い。

Randomized phase II selection designの別の方法として, 患者を2つ以上の試験治療に割り付け, それぞれヒストリカルコントロールと比較するというものもある。群ごとにサンプルサイズ計算をするので, 群の数だけphase II studyをやるのと同じサンプルサイズが必要であるが, 効果と毒性の両方から判断してよい方を次の段階へ進めるといった判断が下せる。これは, 効果と毒性のプロファイルがまったく違う場合に有効と考えられるが, ヒストリカルコントロールとの比較はあまり信用できないため, 選んだ薬剤がphase IIIで既存治療に勝る確率が高いという確証にはつながらない。

2. Reference arm phase II design

Hersonら⁵⁾は, 試験治療群と標準治療群にランダム化するが, サンプルサイズの増大を防ぐため, 直接比較はしないというデザインを提案し

た。標準治療群はヒストリカルコントロールの対象者と試験の対象者が同じ集団かどうか(比較可能かどうか)をチェックするためだけに用い、新治療群はヒストリカルコントロールと比較するというものである。標準治療群の成績がヒストリカルコントロールの成績と変わらなかった場合は、試験治療群とヒストリカルコントロールを比較することに問題ないが、標準治療群が予想より低かったり高かったりした場合はどのように比較するかが問題となる。その場合にはもう一度研究を行えばよいとオリジナルの著者らは言っているが、これは現実的ではないし、もう1回やったからといって比較可能な成績が出るという保証はない。このように、ヒストリカルコントロールと同じかのチェックが失敗すると試験結果が解釈しづらくなるといえる。

3. Phase II/III design

標準治療と新治療でランダム化phase IIをやり、その結果に基づいてそのままphase IIIに進むデザインをrandomized phase II/III designと呼ぶ。いろいろなバリエーションがありえるが、共通していえることとして、同じ対象者を使えるのでサンプルサイズの節約になる一方、phase IIIに進まなかった場合にはphase IIIの準備が無駄になる、といった特徴がある。ここでは、提案されているデザインをいくつか具体的に紹介する。

Schaid⁹⁾は、標準治療Aとそれと比較したい新治療B、さらにはphase II試験をしたい新治療C1, C2, C3...でランダム化比較試験を行い、Cに失敗したらA or Bにさらにランダム化して組み込むというデザインを提案した。このデザインの特徴は、無治療の患者をphase IIに組み込み、失敗したらmore established therapyであるA or Bへ治療を変更するというものである。しかし、無治療の患者をphase IIに入れることに対する倫理的な問題や、前治療としてのCの影響をどう取り扱うかといった問題があり、あまり現実的でないと思われる。

また、Storer⁷⁾は、後継phase IIIと同じランダム化をphase II部分で行うというデザインを提案している。オリジナルの点としては、新治療をヒストリカルコントロールと比較し、よければphase IIIへ、よくなければphase IIとして

終了するというものである。Phase IIの段階ではランダム化した標準治療群との比較を行わないので、ヒストリカルコントロールと比較することの問題点は解消されていない。Phase IIIが終わって、ふたを開けてみたら負けていた、ということも十分ありえる。

Ellenberg⁸⁾は、phase II部分はresponseで比較、少しでも勝っていたらphase IIIへというデザインを提案した。これは、誤ってphase IIで止めてしまう確率が十分低くなるようにして、phase IIIでの検出力低下を防ぐことを意図したデザインである。Phase IIIのサンプルサイズの1/3位、通常のsingle arm phase IIの2倍くらいのサンプルサイズを提案している。このデザインではresponseがprimary endpointである生存期間のよいサロゲートエンドポイントである場合でも、帰無仮説が正しい、すなわち両群に差がない場合にも50%の確率でphase IIIへ進んでしまうので、スクリーニング目的としてはあまり効率がよいとはいえない。

Schaid⁹⁾, Scher¹⁰⁾は、複数の新治療群があっても行うことができるrandomized phase II/III designを提案した。Phase II部分も標準治療群と真のエンドポイントである全生存期間(overall survival; OS)で比較し、ある規準を超えていればその群のみphase IIIへ進むというもので、phase II部分は全体の半分程度のサンプルサイズという提案を行っている。検定の多重性の問題が生じるため α エラーを増大させないためには、なんらかの統計的方法を組み込む必要があるだろう。

私が所属するJCOG (Japan Clinical Oncology Group)でも、実際のいくつかの試験でphase II/III designを採用している。1種類のデザインを用いているのではなく、試験の内容によりさまざまなデザインを採用している。たとえば、phase IIからphase IIIへ進むための判断規準として、

- ・ 新治療群の安全性だけを確認してそれが規準を満たしていればphase IIIへ進む。
- ・ 標準治療群に対し有意に劣っていないことのみ証明できればphase IIIへ進む。
- ・ 両群別々にヒストリカルコントロールと比べて劣っていないかを検証してphase IIIへ進む。

表1 PFSをエンドポイントとした場合のランダム化スクリーニング試験に必要な総イベント数

	検出したいハザード比 Δ (化療 vs. 化療+分子標的薬)			
	$\Delta=1.3$	$\Delta=1.4$	$\Delta=1.5$	$\Delta=1.75$
$(\alpha, \beta) =$ (10%, 10%)	382	232	160	84
$(\alpha, \beta) =$ (10%, 20%)	262	159	110	58
$(\alpha, \beta) =$ (20%, 20%)	165	100	69	36

(文献³⁾より改変)

といったようにphase IIIでの α エラーを消費しない仕組みを組み込んでいる。いますぐにでもphase IIIを行いたい、その前に有効性や安全性の再確認をしておく必要があるといった場合に用いることが多い。そのため、phase II部分で検証的な解析を行うことには重きを置いておらず、phase II部分で α エラーの増大が生じないよう注意を払っているといえる。これらは統計的にかなりの応用問題であり、統計家が関与しても厳密には解決しつくしていない部分が残るが、サンプルサイズの節約とphase IIIを計画する時間の節約がcriticalな場合には、採用する余地があると考えられる。

Phase II/III designをまとめると、メリットとして、①phase II部分のデータをphase IIIでも使える、②phase II終了後phase IIIをすみやかに開始できる、③phase II部分のエンドポイントにはflexibilityがある、④標準治療群をconcurrent controlを使うのでphase II段階での比較の妥当性が上がる、などがあげられる。デメリットとして、①スクリーニングツールとしての性能はよくない、②通常のphase IIに比べてphase II部分のサンプルサイズが大きい、③phase IIIを実施するためのインフラと準備が必要、④phase III部分の検出力を保つためにはphase II部分の α (=controlよりたいしてよくないのに間違っただけでphase IIIに進んでしまう確率)が大きくなる、⑤phase II部分で早く判断できるエンドポイントがなければphase IIの結果が出るまで中断しなければならない、といった点があげられる。したがって、phase II/III designはeffective screeningというより、aggressive interim monitoringを

表2 レスポンスをエンドポイントとした場合のランダム化スクリーニング試験に必要な全体の患者数

	検出したいレスポンスの差 (化療 vs. 化療+分子標的薬)			
	20% vs. 35%	40% vs. 55%	20% vs. 40%	40% vs. 60%
$(\alpha, \beta) =$ (10%, 10%)	256	316	156	182
$(\alpha, \beta) =$ (10%, 20%)	184	224	112	132
$(\alpha, \beta) =$ (20%, 20%)	126	150	78	90

(文献³⁾より改変)

もつphase IIIと考えた方がよいであろう。

4. Randomized phase II screening design

最近提案されたものにランダム化スクリーニングデザイン(randomized phase II screening design)³⁾と呼ばれるものがある。これは、標準治療と新治療を直接に、しかし決定的ではない形(nondefinitive)で比較するものであり、標準治療 vs. 標準治療への分子標的治療のadd onなどといった比較を念頭において提案されたデザインである。これは、ヒストリカルコントロールとの比較による選択バイアスの影響を避けるために標準治療群とランダム化し、かつ、試験の規模は2群で50~100名くらいのサンプルサイズにしたいという欲求を満たすものである。そのためには、治療効果の差に対して α エラー、 β エラーを注意深く設定することによってサンプルサイズ設計を行うことが必要となる。また、決定的な結論を導くことのできるphase III試験ができなくなることがないように、OSとは異なる、より早期のエンドポイントを使うなどといった工夫が必要である。たとえばPFSやresponse、%PFSなどといったものである。ちなみに、ある時点で全員のprogressionを測定できる場合には、その時点での%PFSの方が観察間隔の影響を受けやすいPFSに比べて観察によるバイアスが小さいといえる。

PFSをエンドポイントとした場合のランダム化スクリーニング試験に必要な患者数を表1、responseをエンドポイントとした場合のランダム化スクリーニング試験に必要な患者数を表2に示す。 $\alpha > 20\%$ 、 $\beta > 20\%$ は薦められないので、それ以下で考えても、両群合わせて100名程度あ

るいはそれ以下で設計できる組み合わせがある。つまり、治療効果の差に対して、 α , β を注意深く設定することによって100以下のサンプルサイズ設計が行えるケースがある。たとえば、治療効果をハザード比で表した表1より、 $\Delta > 1.4$ であれば α , β がともに20%の場合必要イベント数は100以下となる。 α が大きすぎるとスクリーニング能力がない、つまり phase II を実施せずに phase I から一足飛びに phase III を実施することと同じになってしまい、逆に β が大きすぎると有効な薬剤の開発を誤って中止してしまうことになる。過度に楽観的な治療効果の差 Δ を考えることも β を大きくすることと同じである。これらから、スクリーニングには $HR=1.5$ (レスポンスでは差が20%) に対して $\alpha=\beta=0.2$ 程度が一つの目安といえるだろう。

ランダム化スクリーニング試験の他の特徴として、同じ薬剤を用いて実施されるスクリーニング試験の数が多くなると false positive の確率が高くなることがあげられる。たとえば、同じ薬剤を用いて異なる3つの対象に試験を実施したとすると、 $\alpha=\beta=0.2$ の場合には、その薬剤が3つの対象すべてに治療効果がある場合に1つも治療効果があることを見出せない確率が5%以下となる一方、3つの対象すべてに治療効果がない場合に少なくとも1つの対象で効果ありと判断してしまう確率は約50%になってしまう。

ランダム化スクリーニングデザインでの重要な点は、得られた結果が definitive phase III から得られたものとは違うことをしっかり意識するという点である。とくに OS をエンドポイントにしているときは、研究者が substantial evidence と思ってしまうことが問題である。そのため、PFS, response rate, %PFS といった、できるだけ OS と異なるエンドポイントを用いることも必要であろう。このようなデザインで実施された試験結果を definitive phase III と考えてはならないのは、何度もランダム化スクリーニング試験を実施する状況では検定の多重性の問題と同様の状況が生じ、 $\alpha=0.05$ が判断の基準として大きすぎるためである。Phase III をやるにはかなりの証拠が必要だが、phase II はそうでもない。したがって、通常の phase II-phase III sequence

と同じような証拠を与えるためには、phase III の中間解析と同じように考えて、たとえば $\alpha < 0.005$ などといった極端な結果が得られた場合にのみ substantial evidence と考え、そのあとの phase III を行う必要がないといった判断の方針を採用するのがよいかもしれない。また、サンプルサイズが小さいことを考えると、有意な結果が得られたとはいっても大きな治療群間差が大きな信頼区間を伴っているはずであり、より精確な推定値、すなわちより精確な治療効果の見積もりを得るためにも、さらなる研究が必要といえる。

具体的に考えてみよう。ランダム化スクリーニングデザインが使えるシナリオとして、

(1) ある製薬企業は、多くの臓器で標準化療に対して add on できるような分子標的薬 X を開発中であるが、薬剤 X は単剤では inactive である。しかし、薬剤 X + 標準化療をヒストリカルコントロールと比較する single arm phase II はあまり信頼できない。このような場合には、複数のスクリーニング試験を行うという戦略が提案できるだろう。それぞれの臓器で実施する一つ一つの試験に対して比較的大きな β エラーを設定しても、どれかの試験では差を検出できる可能性が高まるからである。そのほかのシナリオとして、

(2) 研究者はある特定の疾患の標準化療に add on できるような分子標的薬 Y_1, Y_2, \dots に優先順位をつけなければいけないが、single arm phase II は毒性や feasibility 程度のデータしか与えないし、薬剤の開発状況の違いや製薬企業の reluctance によって、複数の分子標的薬の randomized selection design はやり難い。そこで、複数のスクリーニング試験を行うという戦略が提案できるかもしれない。この場合も、それぞれの試験に対する比較的大きな β エラーが許容できるからである。

まとめ

本稿ではさまざまな randomized phase II design について議論してきた。最後に phase II でどのデザインを選ぶかを簡単にまとめてみたい。

① Response で評価が可能な場合は conventional

phase II design.

- ②複数の試験治療に優先順位をつけたい場合 randomized selection phase II design.
 ③次にphase IIIに進むことがほぼ確実な場合は phase II/III design.
 ④ヒストリカルコントロールとの比較が行いにくい場合は randomized screening design.

とくに新規分子標的薬を標準治療に add on する場合には, definitive phase III を実施可能でかつ研究者が治療効果と α エラーと β エラーをバランスと分別をもって設定できるのなら, ランダム化スクリーニングデザインはよいデザインといえるだろう.

もちろん, デザインが先にあるわけではなく, 治療開発全体のストラテジーの中で適切なデザインを選ぶということが基本である. 最近 randomized phase II が多く議論されているということは, 数少ないが一発逆転狙いの cytotoxic drug の開発という時代から, 単剤での治療効果は必ずしも高くないが, 分子標的薬という数多くの候補薬剤の中から有効なものを選んで開発していくという時代へと世の中が変わっていることを意味している. これは製薬会社や臨床研究者のみならず, 臨床試験に携わる統計家にとっても大きなチャレンジといえる. ぜひ協力してよい治療を開発していきたいと思う.

文 献

- 1) European Organisation for Research and Treatment of Cancer. Phase II trials in the EORTC. The Protocol Review Committee, the Data Center, the Research and Treatment Division, and the New Drug Development Office. *Eur J Cancer* 1997; 33 : 1361.
- 2) Van Glabbeke M, Steward W, Armand JP. Non-randomised phase II trials of drug combinations : Often meaningless, sometimes misleading – Are there alternative strategies? *Eur J Cancer* 2000; 38 : 635.
- 3) Rubinstein LV, Korn EL, Freidlin B, et al. Design issues of randomized phase II trials and a proposal for phase II screening trials. *J Clin Oncol* 2005; 23 : 7199.
- 4) Simon R, Wittes RE, Ellenberg SS. Randomized phase II clinical trials. *Cancer Treat Rep* 1985; 69 : 1375.
- 5) Herson J, Carter SK. Calibrated phase II clinical trials in oncology. *Stat Med* 1986; 5 : 441.
- 6) Schaid DJ, Ingle JN, Wieand S, et al. A design for phase II testing of anticancer agents within a phase III clinical trial. *Control Clin Trials* 1988; 9 : 107.
- 7) Storer BE. A sequential phase II/III trial for binary outcomes. *Stat Med* 1990; 9 : 229.
- 8) Ellenberg SS, Eisenberger MA. An efficient design for phase III studies of combination chemotherapies. *Cancer Treat Rep* 1985; 69 : 1147.
- 9) Schaid DJ, Wieand S, Therneau TM. Optimal two-stage screening designs for survival comparisons. *Biometrika* 1990; 77 : 507.
- 10) Scher HI, Heller G. Picking the winners in a sea of plenty. *Clin Cancer Res* 2002; 8 : 400.

* * *

特集

臨床試験

第I相試験*

山本精一郎**

Key Words : phase I trial, non-cytotoxic drug, cancer study design

臨床試験のデザイン概説

臨床試験はスクリーニングである。薬剤であれば非臨床段階を経て、臨床導入から徐々に安全性や有効性、実施可能性などを調べながら、多くの候補治療がそれぞれの段階でふるい落とされる。最終的に標準治療と決勝戦を行うことによって、新しい標準治療として取って代わる治療を見つけることが臨床試験を実施するもっとも大きな目的である。臨床開発の段階は「相(phase)」としてとらえられることが多いが、これはスクリーニングである治療法の開発段階を便宜的に「呼び名」として、逐次的な複数の相に分けただけである(ICH-E8)¹⁾。慣用としてよく用いられているものの、開発の相という概念は臨床試験の分類の基礎としてふさわしくなく、試験の目的による分類のほうが望ましい。ICH-E8では、臨床試験は、その試験の臨床開発期間における実施時期および目的によって、臨床薬理試験、探索的試験、検証的試験、治療の使用のように分類可能だとしている。Piantadosiの提案する分類も、これに準ずるものである²⁾。つまり、臨床導入から標準治療になるまでには必然的にそれぞれ満たすべき・調べるべき要件があり、

それを調べるのが各試験の目的となる。たとえば、まずは安全性をみながら用量反応研究を行い、推奨用量が決まったら、その用量で探索的に効果を検討し、標準治療以上の効果が期待できそうであれば、次に標準治療とのランダム化比較試験を行う。めでたく標準治療に勝れば、新しい標準治療となる。目的が決まれば、それに応じてデザインは決まってくる。とはいえ、相という概念は慣習的に用いられ、今となってはもともとの「段階」以上に内容を伴う用語として用いられることも多い。本稿の目的は第I相試験に用いられるデザインを紹介することであるが、本筋に立ち戻り、臨床導入第一段階で必要なこと、すなわち安全性を検討して、その後の開発用量を決定するという観点からデザインについて考えてみたい。

第I相臨床試験の標準的な臨床試験デザイン

人間に新しい治療(候補)を行う最初の段階では、それがどのような有害事象(種類と程度)をもたらすかを調べる必要がある。また、どのくらいの治療強度(薬剤の投与量など)がもっとも治療効果をもたらすかも調べる必要がある。効果や有害事象、適した治療スケジュールを調べるためには、薬理作用も調べる必要がある。これらを調べる試験を第I相試験(phase I trial)と

* Study design of phase I clinical trial for non-cytotoxic drug.

** Seichiro YAMAMOTO, Ph.D.: 国立がんセンター がん対策情報センター がん情報・統計部(〒104-0045 東京都中央区築地5-1-1); Center for Cancer Control and Information Services, National Cancer Center, Tokyo 104-0045, JAPAN

呼ぶ。最初にヒトに薬剤を投与する場合だけでなく、複数薬剤の併用のレジメンを検討する場合にも同様の手順が必要となる。単剤の場合、毒性と効果が比例していると考えられるcytotoxic drugの場合には、耐えられる最大の投与量が最大効果をもたらすという考え方から、想定される毒性(用量制限毒性, Dose limiting toxicity: DLT)に対して、最大耐用量(Maximum Tolerated Dose: MTD)を調べ、推奨用量を決定する試験を第I相試験として行うことが一般的である(簡便のために、本稿ではこの考え方を「MTDに基づく用量探索」と呼ぶことにする)(図1)。新治療(候補)をヒトに初めて行うという性質上、薬理作用や有害事象などを逐次モニターし、それをすぐに試験の進捗に反映させる必要があるため、少数の施設(できれば単施設)で行うことが望ましい。

単剤のcytotoxic drugを例に、標準的な第I相臨床試験デザインを紹介する。わが国の抗悪性腫瘍薬の臨床評価方法に関するガイドライン³⁾をはじめ、海外の教科書⁴⁾に紹介されているデザインは、以下のものである。まず、原則として、マウスに対する10%致死量の用量レベル(LD10)の1/10を試験開始用量として各用量を定め、低い方から順にそれぞれの用量を一定数の患者(多くの場合3例)に投与する。もし、ある用量で初めの3例に対しDLTが現れなければ次の用量へ進み、DLTが現れればさらに3例に対し同じ用量を投与する。結果として6人中1人までの毒性発現なら次の用量レベルへ進み、6人中2人以上ならそこで試験を終了し、その用量をMTDとする方法である。増量幅に関する用量レベルの設定はModified Fibonacciの方法が用いられることが多かった。この方法は、6人中2人で終了するため、直感的には33%の毒性が出る用量をMTDとしているように思えるが、実際は33%の毒性が出るより低い用量で止まることが多い。33%発現用量より低い用量でも6人中2人が発現することは十分あるからである。慣習的にこの方法が用いられることが多いが、ストップルール、増量幅ともにとくに科学的な根拠はなく、許容できる毒性の程度を変更できない、試験開始用量が低すぎる場合には効果のまった

くない用量を多くの患者に投与してしまう、といった問題がある。これに対し、まず用量反応曲線を想定し、それを逐次推定した上で、曲線から推定される推奨用量付近(あるいは試験実施者が決定した許容できる毒性に対応する用量付近)を投与し、これを繰り返すBayes流の方法(Continuous Reassessment Method: CRM)が提案され⁵⁾、利用されている場合もあるが、いまだそれほど一般的とはなっていない。Cytotoxic drugの第I相試験で用いられるデザインである3例コホート、CRMともにMTDに基づく用量探索という考え方に基づいている。

分子標的薬の開発に 現在用いられているデザイン

ところが、近年では分子標的薬などのさまざまな作用機序の新薬が開発の中心となってきた。単剤のcytotoxic drugが開発の中心であった時代には、MTDに基づく用量探索が第I相試験の役割であった。MTDが探索できると、MTDあるいはその下の用量を至適用量(optimal dose)として、プライマリエンドポイントを腫瘍縮小効果(たとえば奏効割合)などとした第II相試験を行う。さらに、腫瘍縮小効果としてよい成績を残した薬剤を用いて、標準治療を対照とした第III相試験を行うというのが標準的な開発ステップであった。

これに対し、分子標的薬は毒性と効果が必ずしも比例しないため、第I相試験でMTDに基づく用量探索という考え方をを用いることが最適とはいえない。しかも、これらの薬剤は必ずしも腫瘍縮小効果が高くないため、第II相試験において腫瘍縮小効果をエンドポイントとすることも常に最適とはいえない。このような場合、全生存期間や無増悪生存期間などをエンドポイントとして第II相試験を行うこともあるが、これらのエンドポイントは薬剤の効果だけでなく、試験に登録される患者によって大きく異なってしまふ。つまり、用量選択が適切でなかったり、延命効果のない薬剤であっても、同じ適格規準の中で比較的予後の良い患者を登録することによって、よりよい成績を得てしまう場合がある。恣意的でなくてもこのようなことが生じてしまえ

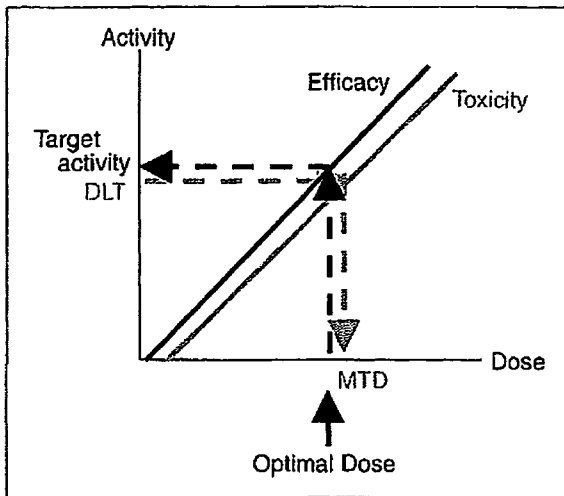


図1 用量-毒性-効果の関係(毒性と効果が比例)

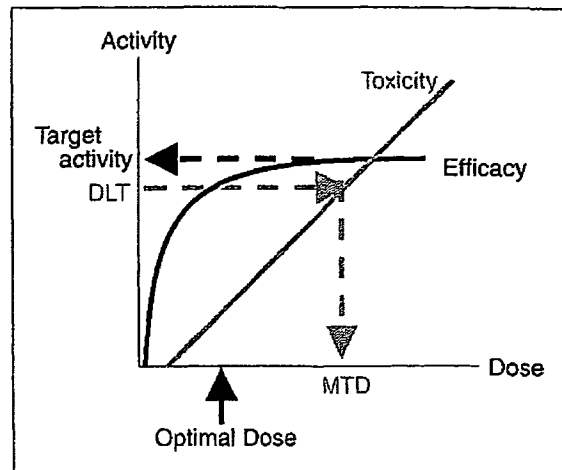


図2 用量-毒性-効果の関係 (MTDより前に効果がプラトー)

ば、その後の第III相試験はnegative resultsに終わってしまう。これらの薬剤に対する第II相以降の臨床試験デザインは別稿に譲るが、臨床試験のデザインを考える上で、開発のすべての段階において、cytotoxic drugとは異なるポイントを考慮する必要がある。

第I相試験のデザインに話を戻そう。分子標的薬の開発のためには、MTDに基づいた用量探索に基礎を置いたこれまでの第I相試験のデザインと同じデザインを用いることができないのか。もしそうなら、どのようなデザインがよいのであろうか。残念ながら、これまでに実現可能性も含めて分子標的薬開発に対するベストなデザインはいまだ提案されていない。NCI sponsored trialのプロトコルをすべてレビューしている米国National Cancer InstituteのClinical Trial Evaluation Program (CTEP)の統計家に尋ねたところ、毒性と効果が必ずしも比例しないような場合でも、大多数の第I相臨床試験で、これまでと同様MTDに基づいた用量探索に基づく3例コホートを用いたデザインが用いられているとのことであった(personal communication)。

そこで、分子標的薬開発において、3例コホートに代表されるMTDに基づく用量探索法に基づくデザインを用いた場合にどのようなことが生じるのか検討してみる。MTDに基づく用量探索法によってOptimal doseが正しく得られるためには以下の条件が満たされていることが必要である⁶⁾。

- 1) 用量が増加すれば、臨床的なベネフィット(効果)が増加する。
- 2) 用量が増加すれば、毒性が増加する。
- 3) 毒性が許容範囲で、効果が得られる用量が存在する。

最近開発が進んでいる分子標的薬では、1)に関して、用量-効果曲線にプラトーがあり、必ずしも高ければ高いほど臨床的なベネフィットが大きくない、あるいは少なくとも用量を上げても少ししか効果が増加しないという場合が考えられる。2)に関して、用量が増加しても必ずしも毒性は増加しない、あるいは毒性が発現するのは十分な効果がみられる以上の用量である場合もある。いくつか具体的に例をあげて考えてみよう。

まず、毒性は用量依存的に増加するが、効果はMTDよりも前にプラトーに達する場合である(図2)。この場合、MTDから推定した用量はsuboptimalといえる。MTDよりも前の時点から効果はプラトーに達しているため、より低い用量のほうが毒性に対する効果のバランスが高いためである。しかし、この場合でもsuboptimalな用量が得られるという意味において、MTD探索デザインはある程度有用であるといつてよいだろう。次に、毒性は用量依存的に増加するが、効果はMTDよりも後にプラトーに達する場合である(図3)。この場合、MTDまでの用量で効果が用量依存的に増加するならMTDに基づく用量

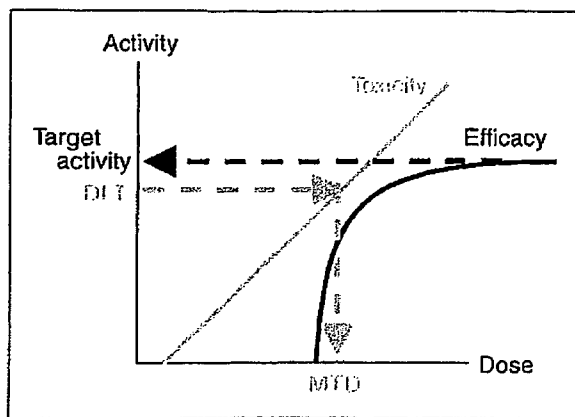


図3 用量-毒性-効果の関係 (MTDより後に効果がプラトー)

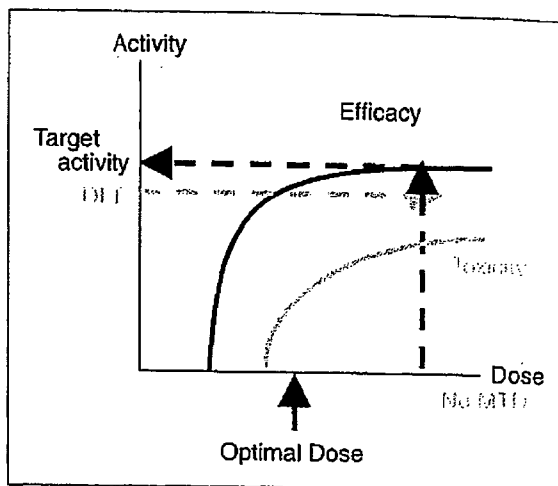


図5 用量-毒性-効果の関係 (毒性がプラトー, 効果もプラトー)

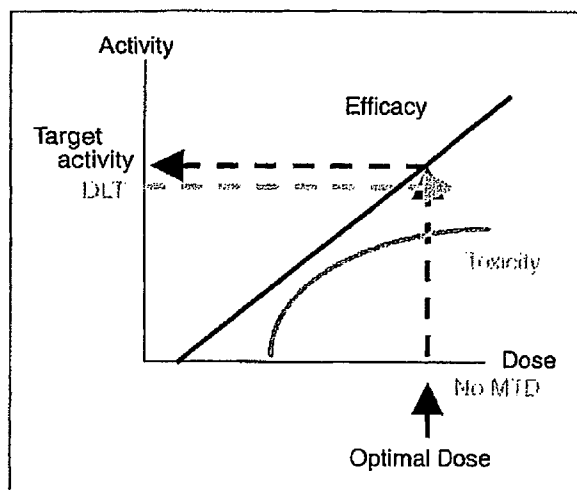


図4 用量-毒性-効果の関係 (毒性がプラトー, 効果が直線的に増加)

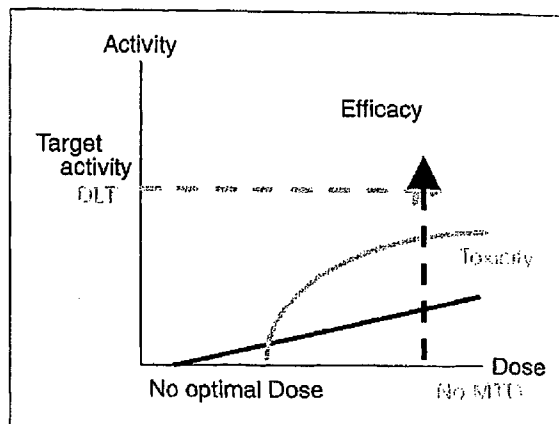


図6 用量-毒性-効果の関係 (毒性がプラトー, 効果が十分でない)

探索法によってoptimal doseを得ることができるが、効果がtarget activityに達しない場合(効果が始まるのがMTDよりも高い用量の場合など)には上記3)の条件を満たさないことになり、optimal doseを見つけることはできない。第I相試験でも効果をモニターできれば効果についてある程度情報が得られるであろう。この場合、よくデザインされた第II相試験を行うことができれば、やはりその結果で開発を中止することになるであろう。

効果が用量依存的に増加し、毒性がプラトーに達するような場合を考えてみよう(図4~6)。この場合、毒性によるMTDが見つけられないので、それ以外の条件で用量増加を中止することになる。効果を理想的にモニターすることがで

ければ、効果がTarget activityに達した用量を見つけることができるが、そうでない場合には、用量-効果曲線がTarget activityを越えて直線的に増加する場合(図4)を除いて、optimalな用量を選択できない。たとえば、用量がtarget activityに達した後プラトーになるような場合には、効果をみないで選択するとsuboptimalな用量を選択することになる(図5)。この場合はまだよいが、効果がTarget activityに達しないような場合には、どのような用量を選んでもoptimalな用量とはいえない(図6)。この場合も、よくデザインされた第II相試験を行うことができれば、その結果で開発を中止することになるであろう。

これらの検討からわかることは、毒性と効果が必ずしも比例関係になくても、ある程度MTD

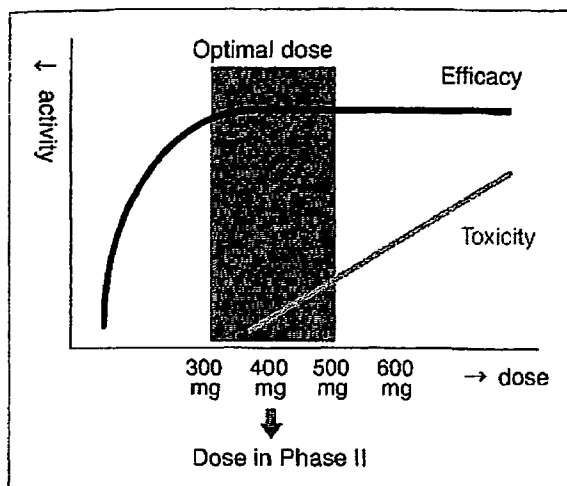


図7 Imatinib trialでの用量-毒性-効果の関係

に基づく用量探索法が有効であることである。しかしながら、いくつかの場合においてうまく働かない場合がある。その場合でも効果をうまくモニターできればoptimalに近い用量を探索することができる。いくつか実例を示そう。図7はimatinib trialの第I相試験における用量-毒性-効果の関係を模式的に表したものである⁷⁾。1,000mgまで用量を上げて毒性はMTDの規準に達せず、効果(Hematologic response)はその用量の範囲内でプラトーに達しており、phase 2では400~600mgを中心として検討が行われた。図8はgefitinibに対する臨床試験であるIDEAL1における用量-毒性-効果の関係を模式的に表したものである⁸⁾。IDEAL1は第I相試験でなく第II相試験である。効果(response)は250mg/dayで18.4%、500mg/dayで19.0%とあまり変わらなかったのに、Grade 3/4の毒性は500mg/dayで多く発現した。結果として(もちろんそれだけが理由ではないだろうが)、日本における申請用量として250mg/dayが用いられた。

しかしながら実際の問題として、各用量に数例しか対象者がいないような第I相試験で効果をモニターするのは難しい。IDEAL1のように毒性プロファイルがある程度わかれば効果も調べられるような第II相試験へ早めに進むか、第I相試験の対象者数をもう少し増やすか、第I相試験のPKやPDなどからtarget inhibitionの情報を得ることによって効果のあたりをつけて複数の候補用量の選択を行う、などいろいろな方法が考

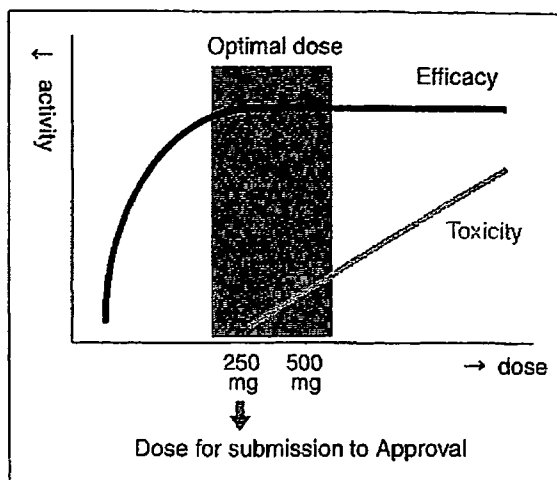


図8 Gefitinib trialでの用量-毒性-効果の関係

えられるが、単純かつ決定的な方法は現在までのところ提案されていない。

まとめに代えて

研究デザインは研究の目的が決まれば、おのずと決まると説明した。まず、第I相、第II相、第III相というデザインがあるというのではなく、治療法開発の段階に従って調べるべき目的があり、それに合わせてデザインをたてることになる。毒性と効果が比例するcytotoxic drugでは比較的単純であった各段階での研究デザインに対し、毒性と効果が必ずしも相関しないといわれている分子標的治療薬の評価はその応用問題といえる。分子標的薬に対して、ベストと考えられる第I相試験のデザインがない現在、基本的にはこれまでに十分経験を積んだMTDに基づく用量探索方法に基づくデザインをもとに、効果を見る工夫を盛り込んだ開発デザインを採用する必要がある。効果と毒性の両方が測定できるような場合には、統計的な定式化によりいくつかのデザインが考えられるが(たとえばHunsburger S. 2005⁹⁾)、それぞれ検証の難しい仮定を含んでおり、たとえ両方がわかる場合にでも、まだまだ実際の適用場面とのすり合わせが必要と思われる。実際には、薬剤と腫瘍の性質によって最適なデザインを頭を絞って考え、結果に対して慎重に検討し、思い切って判断するという時代が当分続くことが予想される。第I相、第II相、第III相それぞれで何をするというのが決まって

いると考えるよりも、ヒトへの導入段階から標準治療とのマッチアップまでをトータルで考えた開発戦略がより有効となるであろう。

文 献

- 1) ICH-E8 臨床試験の一般指針について. ICHガイドラインホームページ. (<http://www.nihs.go.jp/dig/ich/eindex.html>)
- 2) Piantadosi S. Clinical Trials. A Methodologic Perspective. John Wiley & Sons, Inc ; 1997.
- 3) 抗悪性腫瘍薬の臨床評価方法に関するガイドライン. (<http://www.hourei.mhlw.go.jp/hourei/doc/tsuchi/171101-b.pdf>)
- 4) Crowley J. Handbook of statistics in clinical oncology. Marcel Dekker ; 2001.
- 5) Ishizuka N, Ohashi Y. The continuous reassessment method and its applications : a Bayesian methodology for phase I cancer clinical trials. *Statistics in Medicine* 2001 ; 20 : 2661-81.
- 6) Hunsberger S, Rubinstein LV, Dancey J, et al. Dose escalation trial designs based on a molecularly targeted endpoint. *Stat Med* 2005 ; 24 : 2171-81.
- 7) Druker BJ, Talpaz M, Resta DJ, et al. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N Engl J Med* 2001 ; 344 : 1031-7.
- 8) Fukuoka M, Yano S, Giaccone G, et al. A multi-institutional randomized phase II trial of gefitinib for previously treated patients with advanced non-small cell lung cancer (The IDEAL 1 Trial). *J Clin Oncol* 2003 ; 21 : 2237-46.

* * *

Influence of Overweight on Surgical Complications for Gastric Cancer: Results From a Randomized Control Trial Comparing D2 and Extended Para-aortic D3 Lymphadenectomy (JCOG9501)

Toshimasa Tsujinaka, MD,¹ Mitsuru Sasako, MD,² Seiichiro Yamamoto, PhD,³ Takeshi Sano, MD,² Yukinori Kurokawa, MD,³ Atsushi Nashimoto, MD,⁴ Akira Kurita, MD,⁵ Hitoshi Katai, MD,² Toshio Shimizu, MD,⁶ Hiroshi Furukawa, MD,⁷ Satoru Inoue, MD,⁸ Masahiro Hiratsuka, MD,⁹ Taira Kinoshita, MD,¹⁰ Kuniyoshi Arai, MD,¹¹ and Yoshitaka Yamamura, MD,¹² for the Gastric Cancer Surgery Study Group of Japan Clinical Oncology Group

¹Department of Surgery, Osaka National Hospital, 2-1-14 Hoenzaka, Chuo-ku, Osaka 540-0006, Japan

²Gastric Surgery Division, National Cancer Center Hospital, Tokyo, Japan

³Cancer Information and Epidemiology Division, National Cancer Center Research Institute, Tokyo, Japan

⁴Department of Surgery, Niigata Cancer Center Hospital, Niigata, Japan

⁵Department of Surgery, National Shikoku Cancer Center, Matsuyama, Japan

⁶Department of Surgery, International Medical Center of Japan, Tokyo, Japan

⁷Department of Surgery, Sakai Municipal Hospital, Sakai, Japan

⁸Department of Surgery, Tokyo Metropolitan Bokutoh Hospital, Tokyo, Japan

⁹Department of Surgery, Osaka Medical Center for Cancer and Cardiovascular Disease, Osaka, Japan

¹⁰Department of Surgery, National Cancer Center East, Tokyo, Japan

¹¹Department of Surgery, Tokyo Metropolitan Komagome Hospital, Tokyo, Japan

¹²Department of Surgery, Aichi Cancer Center, Aichi, Japan

Background: The impact of overweight on the outcome of gastrectomy with lymphadenectomy is controversial, and data from a well-controlled, randomized study are needed to identify a possible relationship.

Methods: We used data from 523 patients registered for a prospective randomized trial comparing D2 and extended para-aortic D3 lymphadenectomy to compare the effects of body mass index (BMI) and the extent of lymphadenectomy for the development of general or major surgical complications (anastomotic leakage, abdominal abscess, and pancreatic fistula).

Results: Seventy-seven patients were classified as overweight with BMI ≥ 25 , and 38 and 39 of these patients underwent a D2 or D3 lymphadenectomy, respectively. Among the 446 patients classified as nonoverweight with BMI < 25 , 225 received D2 and 221 received D3 lymphadenectomy. Surgical complications, operation time, and blood loss were statistically significantly associated with BMI, and logistic regression analysis revealed that overweight directly affected the occurrence of surgical complications even after considering operation time

Received July 27, 2006; accepted July 27, 2006; published online December 5, 2006.

Members of the Gastric Cancer Surgery Study Group of Japan Clinical Oncology Group are listed in the Acknowledgments.

Address correspondence and reprint requests to: Toshimasa Tsujinaka, MD, E-mail: toshi@onh.go.jp

Published by Springer Science+Business Media, Inc. © 2006 The Society of Surgical Oncology, Inc.

and blood loss as intermediate factors instead of outcome variables. Among patients undergoing D2 lymphadenectomy, being overweight increased the risk for surgical complications and blood loss, whereas overweight was associated with only blood loss and operation time among patients receiving D3 lymphadenectomy.

Conclusions: Overweight increased the risk of surgical complications in patients undergoing gastrectomy both directly and indirectly through operation time and blood loss. The impact of overweight on surgical complications was more evident in patients undergoing a D2 dissection.

Key Words: Overweight—BMI—Complication—Gastric cancer—RCT—JCOG.

The incidence of overweight and obesity has been increasing in the general population, but the impact of overweight on surgical outcomes is unclear. Cancer surgery in overweight patients often takes longer and is associated with greater blood loss than in lean individuals as a result of the presence of excessive fat tissue impairing surgical procedures and lymph node dissection. The influence of overweight on the outcomes, e.g., surgical complications, surgical quality, hospital stay, and prognosis, of gastrectomy with D2 lymph node dissection for patients with gastric cancer is controversial.¹⁻⁵ These data were derived retrospectively from a single institution, but the surgical procedures and disease stages varied.

A prospective study from multiple institutions that use a similar surgical procedure is the ideal means to assess the impact of overweight on surgical outcomes and overall prognosis. A randomized trial, Japan Clinical Oncology Group (JCOG) 9501, was launched in 1995 to explore the potential survival benefit of extended para-aortic D3 dissection over standard D2 dissection. This trial provided the opportunity to prospectively evaluate collected data regarding the effect of overweight on surgical outcome after D2 or D3 dissection. Because a patient's physical condition, including body mass index (BMI), could affect treatment indications for either D2 or D3, an observational study may not correctly compare potential differences between groups. Thus, we used the JCOG data to investigate the interaction of D2/D3 dissection and overweight on surgical complications in a randomized trial. In addition, we examined whether overweight directly influences the occurrence of complications or if the effects of overweight may be mediated by associated factors, such as operation time and blood loss.

PATIENTS AND METHODS

Between June 1995 and April 2001, a total of 523 patients registered in the JCOG9501 study were randomly allocated to either D2 (n = 263) or D3

dissection (n = 260) by balancing the groups according to institution, tumor growth pattern (expansive vs. infiltrative growth) and tumor (T) stage (cT2b vs. cT3/cT4). Patients aged <76 years with histologically proven and resectable primary gastric carcinoma with an estimated depth of SS (invading subserosa: cT2b), SE (penetrating serosa: cT3), or SI (invading adjacent structures: cT4) were recruited after providing informed consent as described elsewhere.⁶ Patients with free cancer cells by cytological examination of peritoneal washes and those with type 4 tumor (linitis plastica type) were excluded.

Patients underwent appropriate gastrectomy with systematic lymphadenectomy as allocated by the study protocol. Perigastric lymph nodes (nodal station nos. 1, 3, 4, 5, and 6 according to the Japanese Classification of Gastric Cancer) and nodes at the base of the left gastric artery (no. 7), along the common hepatic artery (no. 8) and at the base of the splenic artery (no. 11) were routinely resected. Lymph nodes along the hepatoduodenal ligament and behind the pancreatic head (nos. 12 and 13) were resected when the primary lesion was located in the lower third of the stomach. Lymph nodes along the left side of the cardia (no. 2), within the splenogastric ligament (no. 4sa) and at the splenic hilum (no. 10), were resected with the spleen when total or proximal gastrectomy was performed. In patients randomized to a D3 lymphadenectomy group, para-aortic lymph nodes from the level of the celiac trunk down to the root of the inferior mesenteric artery (nos. 16a2 and 16b1) were dissected. The mode of reconstruction after resection was not specified.

Information on complications (including major surgical complications) and patient backgrounds (including height and body weight) was extracted from the case report forms for the trial. In this study, anastomotic leakage, pancreatic fistula, and abdominal abscess are defined as surgical complications. Anastomotic leakage was defined as dehiscence confirmed by radiographic examination that used contrast medium. Pancreatic fistula was diagnosed if

there was prolonged purulent discharge that contained pancreatic juice from the drainage tube. In addition, pneumonia and other complications were evaluated as complications.

According to the World Health Organization classification, BMI ≥ 25 is considered as overweight and BMI < 25 as nonoverweight.⁷ Factors that might affect the risk of overall and major surgical complications, such as sex, age, tumor location, pathological (p) T category (pT2 and pT3 vs. pT4), extent of lymphadenectomy, type of gastrectomy, splenectomy, and pancreatectomy were evaluated as potential confounding factors. The difference in the distribution of these factors between BMI < 25 and BMI ≥ 25 were examined by χ^2 test. The effect of overweight on the complications was evaluated by odds ratio. In addition, the effect of overweight on operating time, amount of blood loss, need for autologous blood transfusion, reoperation, and hospital death was also evaluated by odds ratio. Operating time, blood loss, and the number of retrieved lymph nodes were divided into tertiles as previously described⁸ and used as binary variables by dichotomizing the highest tertiles and the remaining two tertiles because biologically meaningful cutoff points could not be defined. In addition to the univariate analysis, all the analyses were conducted adjusting all the potential confounding factors by logistic regression.

To evaluate the effect of overweight on complications, logistic regression on the complications were conducted with overweight as exposure and operating time and blood loss as intermediate factors in addition to the other potential confounding variables. This analysis reveals whether overweight affects complications directly, or indirectly through these intermediate factors.

To see the difference of the effect of overweight between D2 and D3 dissection, all the analyses were repeated separately for the D2 and D3 subgroups, and these interactions were also evaluated. All statistical analyses were performed SAS software version 8.12 (SAS Institute, Tokyo, Japan). *P* values less than .05 were considered statistically significant, and all tests were two-sided.

RESULTS

Seventy-seven patients were classified as overweight with BMI ≥ 25 , and 38 and 39 of these patients underwent D2 or D3 lymphadenectomy, respectively. In 446 patients classified as nonoverweight with BMI

TABLE 1. Backgrounds of patients according to body mass index (BMI)

Factor	BMI < 25 (n = 446)	BMI ≥ 25 (n = 77)	Total number	<i>P</i> value
Sex				
M	301	57	358	.26
F	145	20	165	
Age				
< 56	137	23	160	.93
56-65	176	31	207	
> 65	133	23	156	
Location				
A (lower third)	188	29	217	.59
M (middle third)	173	33	206	
C (upper third)	85	15	100	
Clinical tumor stage				
cT2b	161	31	192	.38
cT3	268	41	309	
cT4	17	5	22	
Lymph node dissection				
D2	225	38	263	.86
D3	221	39	260	
Type of gastrectomy				
Distal	272	48	320	.82
Total/proximal	174	29	203	
Splenectomy				
No	283	49	332	.98
Yes	163	28	191	
Pancreatectomy				
No	427	74	501	.88
Yes	19	3	22	

< 25 , 225 received D2 and 221 received D3 lymphadenectomy. Total gastrectomy was performed in 199 (38.0%) of 523 patients and proximal gastrectomy in 4; the remaining patients underwent distal gastrectomy. Splenectomy was performed in 191 patients (36.5%) and distal pancreatectomy in 22 (4.2%). The background characteristics of patients with different BMIs are listed in Table 1. There were no statistically significant differences in sex, age, tumor location, clinical T stage, lymph node dissection, type of gastrectomy, and incidence of combined resection between the two groups, and the two groups were well balanced.

In the entire sample, any complications were identified in 128 patients (24.5%), and major surgical complications occurred in 49 patients (9.4%). Among overweight patients, however, the proportion developing either any or surgical complications was 35.1% and 19.5%, respectively. When assessed by univariate analysis, overweight statistically significantly increased the risk for pancreatic fistula, abdominal abscess, operation time, and blood loss (Table 2). Additionally, the number of retrieved lymph nodes was less in overweight patients. Multivariate analysis identified that overweight was significantly associated

TABLE 2. Effect of overweight on postoperative complications and other outcome variables^a

Factors	BMI < 25	BMI ≥ 25	Univariate analysis		Multivariate analysis	
			Odds ratio of BMI > 25 (95% CI)	P value	Odds ratio of BMI > 25 (95% CI)	P value
Operation time (min)						
> 297	141	36	1.90 (1.16–3.10)	.01	2.24 (1.29–3.87)	.004
≤ 297	305	41	–	–	–	–
Blood loss (mL)						
> 710	131	44	3.21 (1.95–5.26)	< .001	3.74 (2.19–6.39)	< .001
≤ 710	315	33	–	–	–	–
Blood transfusion						
Yes	98	17	1.01 (.56–1.80)	.98	1.10 (.59–2.03)	.77
No	348	60	–	–	–	–
No. of retrieved lymph nodes						
≤ 54	137	33	1.69 (1.03–2.77)	.037	1.82 (1.06–3.14)	.031
> 54	309	44	–	–	–	–
Reoperation						
Yes	9	3	1.97 (.52–7.44)	.32	1.85 (.47–7.29)	.38
No	437	74	–	–	–	–
Hospital death						
Yes	3	1	1.94 (.20–18.92)	.56	1.96 (.20–19.50)	.56
No	443	76	–	–	–	–
Any complication						
Yes	101	27	1.84 (1.10–3.10)	.021	1.90 (1.11–3.24)	.019
No	345	50	–	–	–	–
Surgical complication						
Yes	34	15	2.93 (1.51–5.69)	.002	3.35 (1.65–6.78)	< .001
No	412	62	–	–	–	–
Anastomotic leak						
Yes	8	3	2.22 (.58–8.56)	.25	2.14 (.54–8.47)	.28
No	438	74	–	–	–	–
Pancreatic fistula						
Yes	20	10	3.18 (1.43–7.09)	.005	4.18 (1.71–10.22)	.002
No	426	67	–	–	–	–
Abdominal abscess						
Yes	19	10	3.35 (1.50–7.52)	.003	3.51 (1.52–8.12)	.003
No	427	67	–	–	–	–
Pneumonia						
Yes	12	4	1.98 (.62–6.31)	.25	1.88 (.58–6.13)	.29
No	434	73	–	–	–	–
Other complication						
Yes	65	11	0.98 (.49–1.95)	.95	0.97 (.48–1.95)	.93
No	381	66	–	–	–	–

BMI, body mass index; 95% CI, 95% confidence interval.

^a Multivariate covariables: BMI, sex, age, tumor location, clinical tumor stage, lymph node dissection, type of gastrectomy, splenectomy, pancreatectomy.

with pancreatic fistula, abdominal abscess, operation time, and blood loss, and the odds ratios (95% confidence intervals) were 4.18 (1.71–10.22), 3.51 (1.52–8.12), 2.24 (1.29–3.87), and 3.74 (2.19–6.39), respectively. The number of retrieved lymph nodes decreased in overweight patients with an odds ratio of 1.82 (1.06–3.14). When operation time and blood loss were treated as intermediate factors, the odds ratios for the development of pancreatic fistula and abdominal abscess decreased to 3.48 and 2.47, respectively, but were still statistically significant.

We next analyzed the D2 (n = 263) and D3 (n = 260) dissection subgroups (Table 3). In the D2 subgroup, overweight was significantly associated with pancreatic fistula, abdominal abscess, and blood loss

with odds ratios (95% confidence intervals) of 4.74 (1.42–15.89), 4.72 (1.49–14.99), and 2.83 (1.33–6.04), respectively. In the D3 subgroup, only blood loss with an odds ratio of 5.05 (2.27–11.26) and operation time with an odds ratio of 2.27 were significantly associated with overweight, although the interaction P values between the D2 and D3 subgroups were not statistically significant for any of the factors examined.

DISCUSSION

We clearly showed that overweight patients are at increased risk for the development of organ/space

TABLE 3. Effect of overweight on postoperative complications and other outcome variables stratified with lymph node dissection (D2 or D3)^a

Factor	D2 subgroup (n = 263)		D3 subgroup (n = 260)		Interaction P value
	Multivariate odds ratio of BMI ≥ 25 (95% CI)	P value	Multivariate odds ratio of BMI ≥ 25 (95% CI)	P value	
Operation time					
Operation time > 297 min	2.19 (.96–5.02)	.063	2.27 (1.09–4.73)	.028	.95
Blood loss > 710 mL	2.83 (1.33–6.04)	.007	5.05 (2.27–11.26)	< .001	.30
Blood transfusion	1.73 (.70–4.26)	.23	0.78 (.34–1.79)	.56	.20
No. of retrieved lymph nodes ≤ 54	2.73 (1.28–5.85)	.01	1.06 (.43–2.62)	.9	.12
Reoperation	4.21 (.64–27.61)	.13	0.82 (.09–7.39)	.86	.27
Hospital death	6.82 (.40–117.43)	.19	NE	.98	.94
Any complication	2.62 (1.23–5.61)	.013	1.39 (.65–2.98)	.4	.25
Surgical complications	4.20 (1.59–11.10)	.004	2.60 (.91–7.40)	.074	.51
Anastomotic leak	2.77 (.47–16.19)	.26	1.49 (.16–14.09)	.73	.67
Pancreatic fistula	4.74 (1.42–15.89)	.012	3.61 (.96–13.55)	.057	.77
Abdominal abscess	4.72 (1.49–14.99)	.009	2.55 (.73–8.85)	.14	.48
Pneumonia	2.81 (.79–10.04)	.11	NE	.97	.94
Other complications	1.08 (.34–3.37)	.9	0.91 (.37–2.23)	.83	.82

BMI, body mass index; NE, not able to estimate.

^a Covariables: BMI, sex, age, tumor location, clinical tumor stage, type of gastrectomy, splenectomy, pancreatectomy.

surgical site infection (SSI) (abdominal abscess and pancreatic fistula) complications after gastrectomy with D2 or D3 dissection. Risk factors for the development of SSI in abdominal surgery have been intensively investigated. The presence of a preoperative cutaneous abscess or necrosis, sutures or anastomoses of the bowel, postoperative abdominal drainage, surgical treatment for cancer, and postoperative anticoagulant therapy were identified as risk factors for SSI in noncolorectal abdominal surgery.⁹ However, others reported that operation time was the only statistically significant risk factor for SSI after gastrectomy,¹⁰ and in colorectal surgery, diabetes and a 10% weight loss were associated with SSI.¹¹ Among all of these studies, overweight was not identified as a risk factor for SSI. BMI exhibited a direct relationship with operation time in cholecystectomy, colectomy, and unilateral mastectomy, but it was not associated with surgical complications.¹² Thus, BMI may not directly influence the occurrence of surgical complications or SSI in abdominal surgery, but increased operation time and blood loss secondary to BMI may be responsible for any identified negative outcomes. However, we analyzed operation time and blood loss as intermediate factors instead of outcome variables, and BMI was still associated with the development of pancreatic fistula and abdominal abscess, as seen previously.⁸ This fact suggests that BMI has a direct effect on surgical complications besides indirect effects through operation time or blood loss.

Practically, the presence of a large amount of the viscera may disturb drainage of exudates and coag-

ula, and excess fatty tissue may become necrotic more easily as a result of surgical manipulation. In addition, the demarcation between pancreas and fat tissues in overweight individuals is obscure because of greater fat deposition in the pancreas.^{13,14} This could also be relevant in cases of gastrectomy requiring peripancreatic nodal dissection and mobilization of the pancreas. These factors may contribute to the increased occurrence of abdominal abscess and pancreatic fistula in overweight surgical patients.

Whites in general have a higher BMI than Japanese individuals, and the incidence of morbid obesity is marked and growing among patients in the United States and Europe. The proportions of patients with BMI ≥ 25 and BMI > 30 in the present study were only 14.7% and 1.0%, respectively, whereas one-third of the U.S. population is obese (BMI > 27).¹⁵ These differences in patients' physique may partly explain observed differences in mortality and morbidity between the UK Medical Research Council (MRC) and Dutch trials and the present study.^{16,17} The mortality of patients undergoing D2 dissection in the two Western studies was 13% and 10%, whereas morbidity was 46% and 43%. In contrast, we observed only 1.3% mortality and 35.1% morbidity in overweight patients undergoing D2 or D3 dissection. In addition to possible differences in patients' physique, experience and workload volume of surgeons are important factors that could contribute to different surgical outcomes.

In patients undergoing D2, but not D3, dissection, overweight was associated with surgical complications. Although these differences were not statistically

significant, this may be because of low statistical power to test the interactions. In contrast, only the odds ratios of long operation time and excessive blood loss increased were statistically significant in the D3 dissection group, as reported previously.⁶ The increased risk of complications in nonoverweight patients in the D3 subgroup could explain these differences. Indeed, the cumulative incidence of all complications in normal patients was 17.8% in the D2 subgroup and 27.6% in the D3 subgroup. Thus, greater care should be taken in performing gastrectomy not only in all patients undergoing D3 dissection, but also in overweight patients undergoing D2 dissection.

The relationship between overweight and overall prognosis in patients with cancer is an important issue to resolve. The presence of excess fat impairs precise nodal dissection and decreases the yield of lymph nodes. In this study, the number of lymph nodes retrieved from overweight patients was far less compared with nonoverweight patients undergoing a D2, but not D3, dissection. In addition to the quality of lymph node dissection, comorbid conditions associated with overweight, such as cardiovascular diseases, pulmonary dysfunction, diabetes, and hypertension, may negatively affect the prognosis of postoperative patients.¹⁸ The relationship between overweight and overall survival in patients with gastric cancer remains controversial.¹⁻⁴ A conclusive result cannot be obtained without a well-controlled prospective study, and the final results of the JCOG9501 trial should answer this important question. However, the present study provides some insight into this issue.

The proportion of overweight patients in this trial was low (14.7%). Therefore, the obtained results are not definitely conclusive, but they clearly suggest that caution is needed when performing gastrectomy for gastric cancer in overweight patients. In conclusion, overweight increased the risk of surgical complications in patients undergoing gastrectomy with lymphadenectomy.

ACKNOWLEDGMENTS

Supported in part by Grant-in-Aid for Cancer Research from the Ministry of Health and Welfare and the Second Term Comprehensive 10-Year Strategy for Cancer Control by the Ministry of Health and Welfare, Japan. The authors thank Dr. Yoshimura for help with the statistical analysis, Ms. Hongo for data management, and Ms. Sugimoto for secretarial assistance. Participating institutions and

chief participants: National Cancer Center Hospital (M. Sasako, T. Sano), Niigata Cancer Center Hospital (A. Nashimoto, H. Yabuzaki), National Shikoku Cancer Centre (A. Kurita, Y. Kubo), Osaka Medical Center for Cancer and Cardiovascular Diseases (M. Hiratsuka, I. Miyashiro), Osaka National Hospital (K. Fujitani, M. Hirao), National Cancer Centre Hospital East (T. Kinoshita), Tokyo Metropolitan Komagome Hospital (K. Arai, Y. Iwasaki), Aichi Cancer Centre (T. Kito, Y. Yamamura), Osaka Medical College (K. Okajima, M. Tanigawa), International Medical Centre of Japan (O. Kobori, T. Shimizu), Sakai City Hospital, Kanagawa Cancer Centre (H. Furukawa, H. Imamura), Tokyo Metropolitan Bokuto Hospital (M. Kitamura, S. Inoue), Nagaoka Chuo General Hospital (T. Yoshikawa, T. Shimizu), Niigata City General Hospital (K. Aizawa), Cancer Institute Hospital (K. Ota, S. Oyama), Kyoto Second Red Cross Hospital (H. Tokuda, S. Takahashi), Saitama Cancer Centre, Hiroshima City Hospital (Y. Tanaka, K. Uchida), Kanazawa University (K. Miwa, T. Fujimura), Gifu Municipal Hospital (H. Tanemura, H. Oshita), Kagoshima University (T. Aiko, S. Hokita), Iwate Medical University (M. Terashima, K. Saito), and Okayama University (H. Isozaki).

REFERENCES

1. Kodera Y, Ito S, Yamamura Y, et al. Obesity and outcome of distal gastrectomy with D2 lymphadenectomy for carcinoma. *Hepatogastroenterology* 2004; 51:1225-8.
2. Inagawa S, Adachi S, Oda T, Kawamoto T, Koike N, Fukao K. Effect of fat volume on postoperative complications and survival rate after D2 dissection for gastric cancer. *Gastric Cancer* 2000; 3:141-4.
3. Dhar DK, Kubota H, Tachibana M, et al. Body mass index determines the success of lymph node dissection and predicts the outcome of gastric carcinoma patients. *Oncology* 2000; 59:18-23.
4. Barry JM, Blackshaw GRJC, Edwards P, et al. Western body mass indices need not compromise outcomes after modified D2 gastrectomy for carcinoma. *Gastric Cancer* 2003; 6:80-5.
5. Moriwaki Y, Kunisaki C, Kobayashi S, Harada H, Imai S, Kasaoka C. Does body mass index (BMI) influence morbidity and long-term survival in gastric cancer patients after gastrectomy?. *Hepatogastroenterology* 2003; 50:284-8.
6. Sano T, Sasako M, Yamamoto S, et al. Gastric cancer surgery: results of morbidity and mortality of a prospective randomized trial (JCOG9501) comparing D2 and extended para-aortic lymphadenectomy. *J Clin Oncol* 2004; 22:2767-73.
7. Seidell JC, Flegal KM. Assessing obesity: classification and epidemiology. *Br Med Bull* 1997; 53:238-52.
8. Kodera Y, Sasako M, Yamamoto S, Sano T, Nashimoto A, Kurita A. Identification of risk factors for the development of complications following extended and superextended lymphadenectomy for gastric cancer. *Br J Surg* 2005; 92:1103-9.

9. Pessaux P, Msika S, Atalla D, Hay J-M, Flamant Y. Risk factors for postoperative infectious complications in noncolorectal abdominal surgery: a multivariate analysis based on a prospective multicenter study of 4718 patients. *Arch Surg* 2003; 138:314-24.
10. Imai E, Ueda M, Kanao K, Miyaki K, Kubota T, Kitajima M. Surgical site infection surveillance after open gastrectomy and risk factors for surgical site infection. *J Infect Chemother* 2005; 11:141-5.
11. Tang R, Chen HH, Wang YL, et al. Risk factors for surgical site infection after elective resection of the colon and rectum: a single-center prospective study of 2,809 consecutive patients. *Ann Surg* 2001; 234:181-9.
12. Hawn MT, Bian J, Leeth RR, et al. Impact of obesity on resource utilization for general surgery procedures. *Ann Surg* 2005; 241:821-8.
13. Yeo C, Cameron JL, Maher MM, et al. A prospective randomized trial of pancreaticogastrostomy versus pancreaticojejunosotomy after pancreaticoduodenectomy. *Ann Surg* 1995; 222:580-8.
14. Yang YM, Tian XD, Zhuang Y, Wang WM, Wan YL, Huang YT. Risk factors of pancreatic leakage after pancreaticoduodenectomy. *World J Gastroenterol* 2005; 11:2456-61.
15. Kuczmarski RJ, Flegal KM, Campbell SM, Johnson CL. Increasing prevalence of overweight among US adults. *JAMA* 1994; 272:205-11.
16. Cuschieri A, Fayers P, Fielding J, et al. Postoperative morbidity and mortality after D1 and D2 resection for gastric cancer: preliminary results of the MRC randomised controlled surgical trial. *The Surgical Cooperative Group. Lancet* 1996; 347:995-9.
17. Bonenkamp JJ, Songun K, Hermans J, et al. Randomised comparison of morbidity after D1 and D2 dissection for gastric cancer in 996 Dutch patients. *Lancet* 1995; 345:745-8.
18. Pi-Sunyer FX. Medical hazards of obesity. *Ann Intern Med* 1993; 119:655-60.

Risk Factors for Para-aortic Lymph Node Metastasis of Gastric Cancer from a Randomized Controlled Trial of JCOG9501

Eiji Nomura¹, Mitsuru Sasako², Seiichiro Yamamoto³, Takeshi Sano², Toshimasa Tsujinaka⁴, Taira Kinoshita⁵, Hiroshi Furukawa⁶, Toshio Shimizu⁷, Masahiro Hiratsuka⁸, Osamu Kobayashi⁹, Yukinori Kurokawa³ and Nobuhiko Tanigawa¹ on behalf of the Gastric Cancer Surgical Study Group of the Japan Clinical Oncology Group

¹Department of General and Gastroenterological Surgery, Osaka Medical College, Takatsuki, Osaka, ²Gastric Surgery Division, National Cancer Center Hospital, Tokyo, ³Cancer Information and Epidemiology Division, National Cancer Center Research Institute, Tokyo, ⁴Department of Surgery, Osaka National Hospital, Osaka, ⁵Department of Surgery, National Cancer Center Hospital East, Kashiwa, Chiba, ⁶Department of Surgery, Sakai Municipal Hospital, Sakai, Osaka, ⁷International Medical Center of Japan, Tokyo, ⁸Department of Surgery, Itami City Hospital, Itami, Hyogo and ⁹Department of Gastrointestinal Surgery, Kanagawa Cancer Center, Yokohama, Japan

Received May 24, 2006; accepted January 29, 2007

Background: No risk factor has been confirmed for para-aortic lymph node (PAN) metastasis from gastric cancer. To identify the risk factors and the most frequent route of metastasis to PAN, we analyzed the prospective data from a phase III trial.

Methods: In JCOG9501 comparing D2 and D2 + PAN dissection, 260 patients with T2(SS)-T4 gastric cancer underwent radical gastrectomy with PAN dissection. The association between various clinicopathological factors and PAN metastasis was examined.

Results: Macroscopic N stage and tumor size ≥ 5 cm were significant risk factors for PAN metastasis after adjusting for other factors. The proportion of PAN metastasis was clearly different between the N0-1 group and the N2-4 group (2.8% versus 20.5%). In the additional multivariate analysis including 17 regional lymph node stations, station No. 7 was the only station with statistical significance ($P = 0.002$, odds ratio = 41.0).

Conclusion: Macroscopic N stage and tumor size were associated with PAN metastasis, and the lymphatics along the left gastric artery seemed to be the most frequent route to the nodes surrounding the aorta. These findings may be useful in predicting PAN metastasis.

Key words: gastric cancer – para-aortic lymph node metastasis – risk factors – randomized controlled trial

INTRODUCTION

Most of the lymphatic flow from the stomach runs into the para-aortic lymph nodes (PAN), which are located above and below the left renal vein, before it flows into the cisterna chyli lying posterior to the aorta (1,2). Although the lymphatic flow is thought to reach the PAN through several routes (3,4), it is unclear which route is the most frequent access to PAN. In addition, the clinicopathological risk factors to predict metastasis to PAN have been

unknown due to selection bias and contamination in surgically PAN positive patients in the majority of retrospective analyses.

In Japan, the regional lymph nodes of the stomach are generally classified into stations numbered as in Fig. 1 (5). According to the 12th Edition of the Japanese General Rules for the Gastric Cancer Study (GRGCS) (6), the lymph node metastasis of gastric cancer was classified into four categories (N1, N2, N3 and N4) (Table 1, Fig. 1). Metastasis to PAN belonged to N4, while it is grouped into distant metastasis (M1) in the TNM classification (7). In Japan, the extended lymphadenectomy including N1 and N2 categories, known as D2, has been accepted as a

For reprints and all correspondence: Eiji Nomura, Department of General and Gastroenterological Surgery, Osaka Medical College, 2-7 Daigaku-machi, Takatsuki, Osaka 569-8686, Japan. E-mail: sur035@poh.osaka-med.ac.jp