CLINICAL TRIALS ARTICLE

Clinical Trials 2006; 3: 1–9

Estimation of the average causal effect among subgroups defined by post-treatment variables

Yutaka Matsuyama[a] *and Satoshi Morita*[b]

Background In clinical trials, when comparing treatments in a subgroup of patients defined by an event that occurred after randomization is required, the standard estimator that adjusts for the post-treatment variable does not have a causal interpretation.

Purpose To address this problem, we formulate clinically relevant causal estimands using the principal stratification framework developed by Frangakis and Rubin [3], and propose a new estimation method for the principal causal effect.

Methods We consider the comparison of the duration of response among patients who responded to chemotherapy in a cancer clinical trial. Our goal is to estimate the local average treatment effect, that is, the treatment difference among patients who would have responded to either treatment. In order to identify this estimand, we make the assumption that the value of the counterfactual indicator of response is independent of both the actual response status and the outcome variable of interest conditional on the covariates. The proposed estimator is a weighted average of the standard estimators for responders where weights are the probability that the response would have occurred had the patient received the other treatment.

Results The proposed method is applied to data from a randomized phase III clinical trial in patients with advanced non-small-cell lung cancer. The average difference for the duration of response among responders estimated by the proposed method and the standard one was 16.1 (days) and 9.5 (days), respectively. We also evaluate the performance of the proposed method through simulation studies, which showed that the proposed estimator was unbiased, while the standard one was largely biased.

Conclusions We have developed an estimation method for the local average treatment effect. For any type of outcome variables, our estimator can be easily constructed and can be interpreted as the treatment effect among patients who would have had the event in either treatment group. *Clinical Trials* 2006; 3: 1–9. www.SCTjournal.com

Introduction

In clinical trials, it is conceptually difficult to make a treatment comparison in a subgroup of patients defined by an event that occurred after randomization. For example, in cancer clinical trials, one would like to compare the duration of response among patients who responded to chemotherapy. One scientific question behind such analyses is that responses to different treatment regimens will have different response durations, so responders are the primary group of interest. However, such comparison is problematic if the treatment has any effect on response, because the subgroup of responders under the standard treatment and the subgroup of responders under the new treatment are not the

[a]Department of Biostatistics, School of Health Sciences and Nursing, University of Tokyo, Tokyo, Japan
[b]Department of Epidemiology and Healthcare Research, Kyoto University School of Public Health, Kyoto, Japan
Author for correspondence: Yutaka Matsuyama, Department of Biostatistics, School of Health Sciences and Nursing, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: matuyama@epistat.m.u-tokyo.ac.jp

10.1191/1740774506cn135oa

same group of patients. This problem is known to epidemiologists as post-treatment selection bias [1,2], which implies that a comparison of the duration of response among responding subgroups does not have a causal interpretation.

Recently, Frangakis and Rubin [3] have proposed a principal stratification framework with respect to post-treatment variables. The principal stratification approach is a cross-classification of patients defined by the joint potential values of the post-treatment variable under each of the treatments that are being compared. As with the example of the chemotherapy responder described above, patients can be classified into four potential subgroups:

1) those who would respond under either treatment assignment, the true responders;
2) those who would not respond under the standard but would respond under the new treatment, the new treatment only responders;
3) those who would respond under the standard but would not respond under the new treatment, the standard treatment only responders;
4) those who would not respond under either treatment assignment, the non-responders.

The principal causal effect is defined as a comparison of potential outcomes of primary interest within a principal stratum. In a comparison of the duration of response among responders, the treatment difference among patients who would have had a response in either treatment assignment, that is, true responders, is a matter of concern, because the potential values for the duration of response under two treatments are defined only for this subgroup. The key property of the principal stratification is that it is based on the stratification by the baseline potential characteristics of each patient and is not affected by treatment. Rubin [4,5] has called this causal parameter the survivor average causal effect in the context of censored quality-of-life (QOL) data due to death.

In this article, we propose an estimation method for the average causal effect among subgroups defined by a post-treatment variable, the responder average causal effect (RACE). It is important to note that we cannot directly observe the principal stratum to which a patient belongs because the indicator of whether a patient would have responded to the other treatment is a counterfactual variable. Therefore, our approach to this problem is to attempt to predict the probability of a response in each treatment group as a function of covariates and estimate the treatment difference among patients who would have responded to either treatment.

A closely related estimation method for the principal causal effect has been proposed by Gilbert *et al.* [6]. Zhang and Rubin [7] have considered the problem of truncation by death in randomized experiments and derived large sample bounds for the principal causal effect, with or without various identification assumptions. Our approach differs from their approaches in incorporating information from variables related to the post-treatment variable (response) and outcome (duration of response). Furthermore, our approach does not require the assumption which rules out the existence of patients who would respond under control treatment but would not respond under new treatment. This assumption, which is similar to the monotonicity assumption [8], may be reasonable in a placebo controlled study, but is not reasonable in an active controlled study.

Estimation of the average causal effect among responders

The definition of principal causal effects

Consider a randomized clinical trial for cancer treatments with two drug treatment conditions – a standard treatment and a new treatment, and two outcomes: an indicator of response to the drug and the duration of response. We assume that the prerandomization or time-dependent covariates are available. We also assume that there is full compliance and no unintended missing data. The objective is to draw inferences about the effect of treatment on the duration of response. Some patients, however, will not respond to the treatment, with the result that the duration of response is not defined.

We define the potential outcomes of the study patients. Let $\boldsymbol{Z}$ be the vector of treatment assignments for the N randomized patients, with ith element Z_i ($Z_i = 1$ for a new treatment; $Z_i = 0$ for a standard treatment). Let $\boldsymbol{R}(\boldsymbol{Z})$ be the N-vector with ith element $R_i(\boldsymbol{Z})$, which is the indicator of whether the ith patient would respond given $\boldsymbol{Z}$. For patients with $R_i(\boldsymbol{Z}) = 1$, let $Y_i(\boldsymbol{Z})$ be the duration of response given $\boldsymbol{Z}$. In order to limit the possible potential outcomes for each patient, we adopt Rubin's [9] stable unit treatment value assumption (SUTVA) throughout. It states that $R_i(\boldsymbol{Z}) = R_i(\boldsymbol{Z}')$ whenever $Z_i = Z_i'$, and $Y_i(\boldsymbol{Z}) = Y_i(\boldsymbol{Z}')$ whenever $Z_i = Z_i'$ and $R_i(Z_i) = R_i'(Z_i) = 1$. SUTVA implies that potential outcomes for each patient i are unrelated to the assignment $Z_j (j \neq i)$ of other patients, and allows $R_i(\boldsymbol{Z})$ and $Y_i(\boldsymbol{Z})$ to be written as $R_i(Z_i)$ and $Y_i(Z_i)$, respectively. Therefore, under SUTVA, each patient has two potential outcomes for response ($R_i(1)$, $R_i(0)$), and at most two potential outcomes for duration of response ($Y_i(1)$, $Y_i(0)$). For each patient, only one of $R_i(1)$ or $R_i(0)$ is observed. Note that $Y_i(1)\{Y_i(0)\}$ is defined only if $R_i(1) = 1$ $\{R_i(0) = 1\}$. We will also

assume the consistency assumption that, for every individual i, if the actual value of Z_i turns out to be z_i, then the value that R (or Y) would take on if Z_i were z_i is equal to the actual value of R (or Y). This assumption relates the observed outcome to the potential outcomes.

The definition of the average causal effect of treatment on the response is straightforward: $E[R_i(1) - R_i(0)]$, that is, the difference between the average of the response had all patients taken the new treatment and the average of the response had all patients taken the standard treatment [9–11]. This unobservable quantity can be obtained from the observable parameter by $E(R_i|Z_i = 1) - E(R_i|Z_i = 0)$, when the random assignment of treatments is carried out correctly.

Drawing causal inferences about the effect of the treatment on Y is more problematic. A standard method adjusts for the post-treatment variable (R) using the difference between the distributions $E(Y_i|Z_i = 1,\ R_i = 1)$ and $E(Y_i|Z_i = 0,\ R_i = 1)$. This comparison, however, is not a causal parameter, because the two subgroups, $Z_i = 1,\ R_i = 1$ and $Z_i = 0,\ R_i = 1$, will not be comparable unless the event of response is random [1]. To overcome this problem, Frangakis and Rubin [3] have proposed the principal causal effect that is defined as a comparison of potential outcomes under standard versus new treatment within a principal stratum:

$$E[Y_i(1) - Y_i(0) \mid R_i(1) = R_i(0) = 1] \tag{1}$$

As was suggested in Rubin [4,5], the most meaningful inferences about the causal effects on Y can be drawn only for this subgroup, because both $Y_i(1)$ and $Y_i(0)$ are defined only for this subgroup. This population level causal parameter is the effect of the treatment on the duration of response (Y) for a common set of patients, that is, patients who would respond under both treatments. Therefore, this parameter does not suffer from the complications of the standard post-treatment-adjusted one.

Potential values $R_i(1)$ and $R_i(0)$ were also used by Robins and Greenland [2], but, like Rosenbaum [1], they did not use those values to define causal effects adjusted for the post-treatment variable. Instead, they used a framework where both the treatment and the post-treatment variable are controllable, and defined counterfactual values of outcomes Y that would have been observed under assignment to treatment Z and if the post-treatment variable somehow were simultaneously forced to attain a value. In such an approach, the duration of response is "missing" among patients who have not responded, and causal estimands can be defined by comparing the distribution between the randomized groups. This framework is not compatible with the studies we consider, which do not directly control the post-treatment variable. Specifically, the duration of response among patients who have not responded to chemotherapy is not really missing data, which would imply a hidden value, but is nonexistent and is simply undefined.

Proposed estimation method

The primary estimand of interest is the difference in outcome Y in the group of patients that would have responded under both treatments, defined by Equation (1). This local average treatment effect can be written as follows:

$$\mu = \frac{E[\{Y_i(1) - Y_i(0)\}R_i(0)R_i(1)]}{E[R_i(0)R_i(1)]} \tag{2}$$

The quantity $R_i(0)R_i(1)$ in both the numerator and denominator of Equation (2) takes the value of one for any patient who would have responded under both treatments and takes the value of zero for all other patients. It is not possible to estimate Equation (2) without introducing assumptions, because the joint distributions involved in the numerator and denominator of Equation (2) are not observable. For example, when the treatment has no effects on response, that is, $R_i(z) = R_i(1 - z)$, we can estimate Equation (2) from the observed data as

$$\frac{\sum_j Y_j(1)R_j(1)}{\sum_j R_j(1)} - \frac{\sum_k Y_k(0)R_k(0)}{\sum_k R_k(0)} \tag{3}$$

where j indexes over patients assigned to group $z = 1$ and k indexes over patients assigned to group $z = 0$. However, if the treatment has any effect on response, the estimator Equation (3) from the observed responders will in general be a biased estimate of the causal parameter owing to the post-treatment selection bias.

In order to identify Equation (2), we make the following assumption about the potential outcomes:

$$\Pr[R_i(1-z) = 1 \mid R_i(z), Y_i(z), X_i] = \Pr[R_i(1-z) = 1 \mid X_i] \tag{4}$$

where X_i represents the prerandomization or time-dependent covariates. This assumption means that the probability that the response would have been observed had the patient received the other treatment can be explained only by measured covariates X_i. Let $w_i(z) = E[R_i(z)|X_i]$ be the expected value of $R_i(z)$ conditional on X_i for $z = 0,1$. Then, under the assumption of Equation (4), we have

$$\begin{aligned} &E[Y_i(z)R_i(z)R_i(1-z) \mid X_i] \\ &= E[Y_i(z)R_i(z) \mid X_i]E[R_i(1-z) \mid X_i] \\ &= E[Y_i(z)R_i(z) \mid X_i]w_i(1-z) \\ &= E[Y_i(z)R_i(z)w_i(1-z) \mid X_i] \end{aligned}$$

From this equation, conditional on X_i and with a consistent estimator $\hat{w}_i(z)$ of $w_i(z)$ for $z = 0,1$, an estimator of Equation (2) is given by

$$\hat{\mu} = \frac{\sum_j Y_j(1)R_j(1)\hat{w}_j(0)}{\sum_j R_j(1)\hat{w}_j(0)} - \frac{\sum_k Y_k(0)R_k(0)\hat{w}_k(1)}{\sum_k R_k(0)\hat{w}_k(1)} \quad (5)$$

where j indexes over patients assigned to group $z = 1$ and k indexes over patients assigned to group $z = 0$.

Although the probabilities of response under the other treatment are unknown, we can predict them from the data in each treatment group. Therefore, our proposed estimation procedure for Equation (2) consists of the following three steps:

1) Modelling: A model such as logistic regression is used to predict the probability of response in each treatment group as a function of covariates.
2) Prediction: Using the estimates of the regression parameters in the other treatment group estimated in step (1), the probability that the response would have occurred had the patient received the other treatment is predicted in each patient.
3) Weighting: The usual analysis comparing the duration of response between treatment groups (ie, estimation of the difference in means) is conducted among the observed responders using the individual-specific weight, which is the estimated probability in step (2).

We provide an alternative explanation of the above step (3). We assume the following causal model for the potential outcomes on principal stratum of the true responders:

$$E[Y_i(z) \mid R_i(1) = R_i(0) = 1] = \beta_z \quad (6)$$

where β_z ($z = 0,1$) is the mean duration of response in each treatment group. We contrast the causal model (6) with the following association model for the observed responders:

$$E[Y_i \mid R_i = 1, Z_i] = \beta_0'(1 - Z_i) + \beta_1' Z_i \quad (7)$$

Assuming no other bias such as measurement error, we can unbiasedly estimate the associational parameter $\beta' = (\beta_0', \beta_1')$ by fitting Equation (7) to the observed data. If there is no post-treatment selection bias, the parameters of models (6) and (7) are equal. As a consequence, associational estimate of β' is also an unbiased estimate of the causal parameter $\beta = (\beta_0, \beta_1)$. If the treatment has any effect on response, then $\beta \neq \beta'$ and unweighted estimate of β' is a biased estimate of the causal parameter β owing to the post-treatment selection bias. However, even when event of response is not random, if the assumption (4) is true, one can obtain unbiased estimate of the causal parameter β of model (6) by fitting the association model (7) using individual-specific weights $\hat{\Pr}(R_i(1 - Z_i) = 1|X_i)$. Again, in practice, individual-specific weight is unknown and one must estimate it from the data by specifying a model.

The use of estimated weights induces within-individual correlation, which invalidates the model-based standard error estimates outputted by many standard statistical packages. Two methods are used to construct the confidence intervals for the weighted estimator. The first is based on a robust variance estimate [12,13]. The robust variance estimator provides conservative confidence intervals for the parameter of interest θ, that is, the 95% Wald confidence intervals calculated as $\theta \pm 1.96 \times$ (robust standard error) is guaranteed to cover the true value θ at least 95% of the time in large samples [14,15]. The robust intervals are conservative because they do not account for the fact that the weights are estimated, and estimating the weights shrinks the variance of our weighted estimator. The second method is based on robust variance that accounts for the variability of the weights (see Appendix). The observation that a weighted estimator that uses the estimated weights has smaller variance than one that uses the true weights has been discussed by Robins *et al.* [16] and a series of papers by Rubin and Thomas [17–19].

Simulation studies

To evaluate the performance of the proposed estimation method, we carried out simulation studies. We simulated data from two treatment groups, coded as $z = 0$ (standard treatment) or $z = 1$ (new treatment). The simulations were based on 1000 replications, so that the estimated coverage probability of a true 95% confidence interval would have a simulation accuracy of approximately 1.35%. For each subject i, a potential outcome variable under the assigned group z, $Y_i(z)$, was generated via the linear model, $Y_i(z|x_i, \varepsilon_i) = \beta_{0,z} + x_i + \varepsilon_i$, where $(\beta_{0,0}, \beta_{0,1}) = (50,60)$. A covariate X_i was generated from a normal distribution with a mean of 10 and a standard deviation of 20. The random error ϵ_i was generated from a normal distribution with a mean of zero and a standard deviation of 5. For each subject i, a potential outcome variable under the other treatment, $Y_i(1 - z)$, was also generated from the above model.

For each subject i, a potential response indicator $R_i(0)$ under the standard treatment was generated via the logistic regression model

$$\Pr(R_i(0) = 1 \mid x_i) = \frac{\exp(\alpha_{0,0} + \alpha_1 x_i)}{1 + \exp(\alpha_{0,0} + \alpha_1 x_i)} \quad (8)$$

In model (8), $\alpha_1 = \log(1.07)$, so that 10-fold increases in X_i produces two-fold increases in the odds of response. $\alpha_{0,0} = -0.08, -0.69, -1.28, -1.93$ and -2.70, so that the probability of response under the standard treatment was nearly 60%, 50%, 40%, 30% and 20%, respectively. For the potential response under the new treatment, the following logistic regression model was assumed

$$\Pr(R_i(1) = 1 \mid x_i, y_i(0) r_i(0)) = \frac{\exp(\alpha_{0,1} + \alpha_1 x_i + \alpha_2 y_i(0) r_i(0))}{1 + \exp(\alpha_{0,1} + \alpha_1 x_i + \alpha_2 y_i(0) r_i(0))} \quad (9)$$

$\alpha_{0,1} = 1.35, 0.58, -0.08, -0.69$ and -1.28, so that the probability of response under the new treatment was nearly 80%, 70%, 60%, 50% and 40% when $\alpha_2 = 0$, respectively. In model (9), two situations were considered: $\alpha_2 = 0$ (situation A), corresponding to the assumption (4), and $\alpha_2 \neq 0$ (situation B), corresponding to a departure from the assumption (4). The selection bias parameter α_2 can be interpreted as the conditional log-odds ratio for response under the new treatment between subjects who differ by 1 in $y_i(0)r_i(0)$. In situation B, $\alpha_2 = 0.1$, which implies that, after conditioning on a covariate X_i, a 10-fold increase in the duration of response under the standard treatment produces an $\exp(1) = 2.72$ times increased odds of response under the new treatment. We are interested in the situation $\alpha_2 > 0$, because among subjects with the same covariate X, those with longer duration of response might be more likely to respond under the other treatment than those with shorter duration of response.

An equal sample size of 100 for each group was randomly generated (total sample size was 200). In situation A, we compared the results from the standard method whose analysis model is given in Equation (7) with those from the proposed method in which weights were estimated by the logistic regression model that included X_i as a covariate. In situation B, only the proposed method was applied. For both situations, the result from the subjects who had a response under both treatments, that is, a true responder stratum, was regarded as a true value in each replication. Therefore, the average true treatment effect for duration of response was $10(=\beta_{0,1} - \beta_{0,0} = 60 - 50)$.

Table 1 shows the results for the case of $\alpha_2 = 0$ in model (9). Each row of Table 1 reports the Monte Carlo mean bias, mean squared error (MSE), and coverage probability of the nominal 95% large sample confidence intervals, for the estimate of difference in Y_i according to the combinations of the response rate in each group. Examining rows 1–5 of Table 1, one can see that under the assumption (4) the proposed estimator is nearly unbiased, while the standard one is largely biased. For both methods, MSE is increasing with the decrease of the numbers of responders. The coverage probabilities for the proposed estimator based on the robust variance, which accounts for the variability of the weights, are close to the nominal level of 95%. One can also see that ignoring the fact that the weights are estimated leads to conservative coverage probability in all situations. Under-coverage rate in row 5 is due to small sample sizes of responders. The bias of the standard estimator is also reflected in the smaller coverage probabilities. Examining rows 6–7 of Table 1, we observe that both methods give unbiased estimators, as expected, when the response rates are equal between the treatment groups.

Table 2 shows the results for the case of $\alpha_2 = 0.1$ in model (9). Examining rows 6–7 of Table 2, we observe that, even in situation B, the proposed estimator is unbiased when the response rates are equal between the treatment groups. The coverage probabilities based on the robust variance, which accounts for the variability of the weights, are close to the nominal level of 95%. However, examining rows 1–5 of Table 2, one can see that the proposed estimator is slightly biased when the assumption (4) is not satisfied. The small degree of bias was due to the fact that, in the above simulation, the weight

Table 1 Simulation results (Ignorable missingness)

	Response rate		Proposed method				Standard method		
No	New	Standard	Bias	MSE	95% coverage[a]	95% coverage[b]	Bias	MSE	95% coverage
1	80%	60%	0.01	11.87	94.5%	96.2%	−4.01	29.88	87.6%
2	70%	50%	0.03	14.74	95.6%	96.8%	−4.21	36.15	83.8%
3	60%	40%	−0.05	18.50	95.8%	96.7%	−4.35	41.43	82.9%
4	50%	30%	−0.12	29.35	95.3%	95.5%	−4.97	61.10	81.0%
5	40%	20%	−0.01	52.92	92.6%	93.4%	−5.57	94.55	81.0%
6	60%	60%	−0.06	13.90	95.3%	96.8%	0.01	16.17	97.1%
7	40%	40%	−0.15	25.63	95.2%	96.7%	−0.18	31.54	96.4%

[a]Robust 95% confidence intervals, which account for variability of weights.
[b]Robust 95% confidence intervals, which do not account for variability of weights.

Table 2 Simulation results (Non-ignorable missingness)

No	α_2	Response rate		Proposed method			
		New	Standard	Bias	MSE	95% coverage[a]	95% coverage[b]
1	0.1	80%	60%	0.35	12.45	98.3%	98.9%
2		70%	50%	0.84	16.32	97.6%	98.2%
3		60%	40%	1.02	20.25	98.2%	98.6%
4		50%	30%	1.36	32.68	96.5%	97.2%
5		40%	20%	0.98	42.36	94.8%	95.6%
6		60%	60%	0.05	15.41	97.1%	97.6%
7		40%	40%	−0.07	23.51	97.8%	98.2%

[a]Robust 95% confidence intervals, which account for variability of weights.
[b]Robust 95% confidence intervals, which do not account for variability of weights.

prediction model was misspecified in only standard treatment group. The bias of our estimator will, in general, be larger as the weight prediction model is mis-specified in both groups. To reduce this bias of our estimator, it is important to collect covariates which satisfy the assumption (4), and to include these covariates as explanatory variables in the weight prediction model.

Application to advanced non-small-cell lung cancer trial data

Our illustrative application is a randomized phase III clinical trial of advanced non-small-cell Lung cancer (NSCLC) Full details on the design, conduct, and the main clinical results have been reported [20]. A total of 398 patients with previously untreated NSCLC were randomized to receive Cisplatin + Irinotecan (CPT-P), Cisplatin + Vindesine (VDS-P), or Irinotecan alone in order to compare the survival rate. Here we used the first two treatment groups to compare the duration of response among responders. The response rate was 43.7% (55/126) for patients in the CPT-P group and 31.7% (38/120) for those in the VDS-P group. The difference in response rates was 12.0% with the 95% confidence interval (−0.03%, 24.0%).

We used a logistic regression model to predict the probability of response in each treatment group. Sex, age, stage and performance status were used as the covariates. Figure 1 shows the distribution of estimates of weight in each group, which is the probability of response that would have occurred had the patient received the other treatment.

Table 3 shows the average duration of response among responders estimated by the proposed method and the standard one. The average difference for the duration of response estimated by the standard method was 9.5 (days), which would be biased because of the difference in response rates between the treatment groups. The estimated difference by the proposed method was 16.1 (days) with the robust 95% confidence interval (−14.6,46.7). The narrower intervals were obtained by accounting for the variability of the weights.

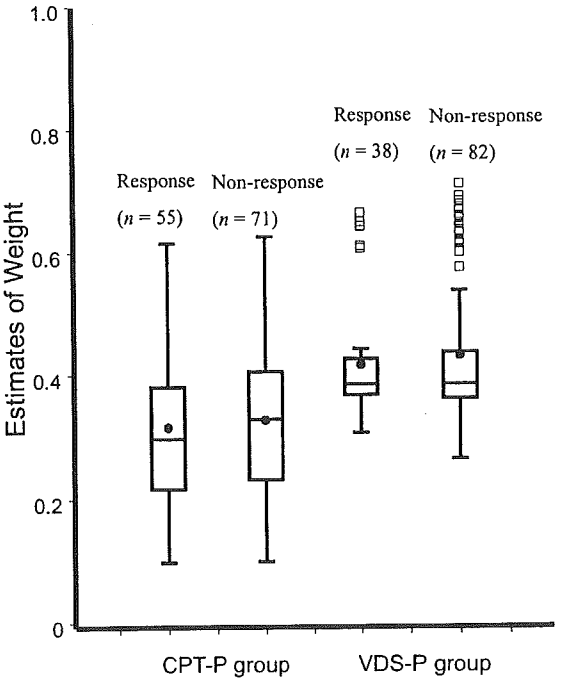

Figure 1 Distribution of weights (the probability of response that would have occurred had the patient received the other treatment) according to the response status in each treatment group. Each box shows the location of the mean (●), median (middle horizontal bar), and quartiles (border horizontal bars). Vertical lines extend to the most extreme observations which are no more than 1.5 × IQR (interquartile range) beyond the quartiles. Observations beyond the vertical lines are plotted individually (□)

Table 3 Estimated average duration of response (days) among responders in NSCLC data

	Proposed method				Standard method		
Group	Mean	Difference	95% CI[a]	95% CI[b]	Mean	Difference	95% CI
CPT-P	89.7	16.1	−14.6, 46.7	−15.9, 48.0	86.6	9.5	−20.7, 39.7
VDS-P	73.6				77.1		

[a]Robust 95% confidence intervals for the difference, which account for variability of weights.
[b]Robust 95% confidence intervals for the difference, which do not account for variability of weights.

Our proposed approach can also estimate the size of the four potential subgroups (true responders, CPT-P only responders, VDS-P only responders and non-responders). For example, the number of true responders can be estimated by summing the estimates of weight among responders which is shown in Figure 1. The number of CPT-P only responders can be estimated by summing both the estimates of (1−weight) among responders in the CPT-P group and the estimates of weight among non-responders in the VDS-P group. Table 4 shows these results. In the NSCLC data, 13.5% of the patients were true responders, while 38.3% of the patients were non-responders who would not respond under either treatment assignment. For the CPT-P only responders, who comprised 29.7% of the population, there is a distribution of Y (duration of response) in the CPT-P assignment group. The average duration of response in the CPT-P group for patients who would not respond under the VDS-P treatment, $E[Y_i|Z_i = 1, R_i = 1, R_i(0) = 0]$, was estimated by weighted analysis in which weights are the probability that the response would not have occurred had the patient received the VDS-P treatment. This resulted in a point estimate of 85.1 (days) with the 95% confidence interval (69.0,101.3). Similarly, the estimated average duration of response in the VDS-P group for patients who would not respond to the CPT-P treatment, $E[Y_i|Z_i = 0, R_i = 1, R_i(1) = 0]$, was 79.5 (days) with the 95% confidence interval (53.8,105.3).

Table 4 Estimated size of the four potential subgroups in NSCLC data

		VDS-P		
		Response	Non-response	Total
CPT-P	Response	33.2 (13.5%)	73.1 (29.7%)	106.3
	Non-response	45.4 (18.5%)	94.3 (38.3%)	139.7
		78.6	167.4	246

Discussion

Within the framework of Frangakis and Rubin [3], this paper develops methods for causal inference in the always-responded principal stratum, that is, we have proposed an estimation method for the "local" average treatment effect, which is the average effect of treatment among responders. Alternatively, causal inference could be made using a missing data framework that assumes all randomized subjects will respond, and thus will have the duration of response. In such an approach, the duration of response is treated as missing in patients who have not responded, and causal estimands can be defined based on the contrasts of the duration of response distributions for the standard and new treatment groups. The goal of assessing such estimands is to compare outcome distribution between the randomized groups had all patients responded during the study. Robins *et al.* [21] developed an IPCW (inverse probability of censoring weighted) method that could be used for causal inference of this estimand, which is the "global" average treatment effect in the entire study population. The drawback of such missing data approach for the present application is that the causal estimand may not be relevant or interpretable, because it is unrealistic to suppose that all patients would respond. Frangakis and Rubin [3] criticize use of such a causal estimand because it uses nonexistent "*a priori*" counterfactuals. However, there are similarities between our estimation method for outcomes in a principal stratum and the IPCW method. The underlying idea for both methods is to base estimation on the observed outcomes but weight them by some quantities for predicting each target population. The IPCW weight is the inverse of the probability of remaining in the study under the MAR (missing at random) assumption. While, under the ignorability assumption (4), our approach weights the contribution of the observed responders by their probability of response under the other treatment.

The proposed method can be used in a variety of applications where the comparison of treatments adjusted for post-treatment variables are required.

As analysed in this paper, the comparison of outcomes among patients who responded to treatment is one example. Gilbert *et al.* [6] considered a randomized study to evaluate the efficacy of a preventive HIV vaccine among infected subjects. They used a similar framework of potential outcomes to formulate the causal estimands, which were defined in terms of the distributions of potential viral loads given assignment to receive vaccine or placebo for subjects in the always-infected principal stratum. Another example is the analysis of QOL data with censoring due to death [4,5,7]. Standard approaches for missing data attempt to estimate the treatment effect that would have been observed if (contrary to fact) all patients had continued to be observed until the end of the study [21], which is the global average treatment effect. This analysis will be reasonable if the post-treatment variable is controllable. However, QOL data for censored cases due to death is "missing" because a null value exists [3]. Therefore, it is reasonable to think that estimation of the causal effects of treatment on QOL is restricted by the life of patients, which can be affected by the treatment and is a post-treatment variable.

Our estimation method is simple in that it requires only the prediction of the probability of the event in each treatment as a function of covariates. For any type of outcome variables, the proposed weighted estimator can be easily constructed among patients for whom the event has occurred and can be interpreted as the treatment effect among patients who would have had the event in either treatment group. This weighted analysis can be easily fitted in many standard statistical packages. However, the correctness of the causal inferences from our weighted analysis depends on the key assumption (4) that the value of the counterfactual indicator of response is independent of both the actual response status and the outcome variable of interest conditional on the covariates. This residual independence between the outcome variables and the counterfactual response is a non-identifiable assumption and is not testable from the observed data. If the covariates which satisfy the conditional independence assumption (4) are available and these covariates are entered as explanatory variables in the weight prediction model, the selection bias due to these unmeasured individual characteristics, that is, the departures from assumption (4), will be small. For example, when the time-dependent covariates or another outcome such as side effects are recorded and are related to response in addition to the baseline covariates, all these factors should be included in the weight prediction model. Therefore, for this issue, it is important to collect data on a sufficient number of covariates for the outcome variables to ensure that the assumption of no unmeasured covariates will be at least approximately true. In our NSCLC data, sex, age, stage and performance status, which were most clinically important factors for tumour response, were used as the covariates to predict the probability of response in each group. The subject matter experts agreed that these covariates used in the analysis could be presumed to be most predictive of tumor response and response duration under either treatment group. Therefore, the assumption (4) seems to be reasonable in our example. Otherwise, it will be necessary to extend the proposed method to investigate the sensitivity of our inferences to the fundamental assumption of no unmeasured covariates [22]. Alternative approach to sensitivity analysis as means for summarizing uncertainty due to unmeasured factors is to compute bounds for the causal effect [7,23]. Zhang and Rubin [7] have derived such bounds for the principal causal effects (Equation 2), although the bounds are often so wide as to be useless for making decisions.

Acknowledgements

We are grateful to the West CPT-11 Lung Cancer Study Group (Dr H Niitani, Dr M Fukuoka and other investigators) for providing us with these valuable data. We would like to thank Tosiya Sato, PhD for his helpful comments on this work. We also thank the referee, the associate editor, and the editor for their comments, which led to a much improved version of the paper. This research was supported in part by a Grant-in-Aid for Scientific Research (A) No. 16200022.

References

1. **Rosenbaum PR.** The consequences of adjustment for a concomitant variable that has been affected by the treatment. *J R Stat Soc Series A - General* 1984; **147**: 656–66.
2. **Robins JM, Greenland S.** Identifiability and exchangeability of direct and indirect effects. *Epidemiology* 1992; **3**: 143–55.
3. **Frangakis CE, Rubin DB.** Principal stratification in causal inference. *Biometrics* 2002; **58**: 21–29.
4. **Rubin DB.** More powerful randomization-based *p*-values in double-blind trials with non-compliance. *Stat Med* 1998; **17**: 371–85.
5. **Rubin DB.** Comment on "Causal inference without counterfactuals", by A. P. Dawid. *J Am Stat Assoc* 2000; **95**: 435–37.
6. **Gilbert PB, Bosch RJ, Hudgens MG.** Sensitivity analysis for the assessment of causal vaccine effects on viral load in HIV vaccine trials. *Biometrics* 2003; **59**: 531–41.
7. **Zhang JL, Rubin DB.** Estimation of causal effects via principal stratification when some outcomes are truncated by "death". *J Educ Behave Stat* 2003; **28**: 353–68.

8. **Angrist JD, Imbens GW, Rubin DB.** Identification of causal effects using instrumental variables (with discussion). *J Am Stat Assoc* 1996; **91**: 444–72.
9. **Rubin DB.** Bayesian inference for causal effects: the role of randomization. *Annals Stat* 1978; **6**: 34–58.
10. **Rubin DB.** Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol* 1974; **66**: 688–701.
11. **Holland PW.** Statistics and causal inference (with discussion). *J Am Stat Assoc* 1986; **81**: 945–70.
12. **Huber PJ.** The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the Fifth Berkley Symposium in Mathematical Statistics and Probability.* University of California Press, 1976: 221–33.
13. **Liang KY, Zeger SL.** Longitudinal data analysis using generalized linear models. *Biometrika* 1986; **73**: 13-22.
14. **Robins JM.** Marginal structural models versus structural nested models as tools for causal inference. In Halloran ME, Berry D eds. *Statistical models in epidemiology: the environment and clinical trials.* Springer-Verlag, 1999: 95–134.
15. **Robins JM, Hernan MA, Brumback B.** Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000; **11**: 550–60.
16. **Robins JM, Rotnitzky A, Zhao LP.** Estimation of regression coefficients when some regressors are not always observed. *J Am Stat Assoc* 1994; **89**: 846–66.
17. **Rubin DB, Thomas N.** Affinely invariant matching methods with ellipsoidal distributions. *Annals Stat* 1992; **20**: 1079–93.
18. **Rubin DB, Thomas N.** Characterizing the effect of matching using linear propensity score methods with normal distributions. *Biometrika* 1992; **79**: 797–809.
19. **Rubin DB, Thomas N.** Matching using estimated propensity scores: Relating theory to practice. *Biometrics* 1996; **52**: 249–64.
20. **Negoro S, Masuda N, Takada Y** *et al.* Randomised phase III trial of irinotecan combined with cisplatin for advanced non-small cell lung cancer. *Br J Cancer* 2003; **88**: 335–41.
21. **Robins JM, Rotnitzky A, Zhao LP.** Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *J Am Stat Assoc* 1995; **90**: 106–21.
22. **Robins JM, Rotnitzky A, Scharfstein DO.** Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models. In: Halloran ME, Berry D eds. *Statistical models in epidemiology, the environment, and clinical trials.* Springer, 1999: 1–94.
23. **Balke A, Pearl J.** Bounds on treatment effects from studies with imperfect compliance. *J Am Stat Assoc* 1997; **92**: 1171–78.

Appendix

The causal parameter $\beta = (\beta_0, \beta_1)$ of model (6) can be obtained from the solution to the following weighted estimating equations among responders $U(\beta, \hat{\alpha}) = \Sigma d(Z_i; \beta) w_i(\hat{\alpha}_{1-z})[Y_i - g(Z_i; \beta)] = \Sigma U_i(\beta, \hat{\alpha}) = 0$, where $g(Z_i; \beta) = \beta_0(1 - Z_i) + \beta_1 Z_i$, $d(Z_i; \beta) = \partial g(Z_i; \beta)/\partial\beta$, and $\hat{\alpha} = (\hat{\alpha}_0, \hat{\alpha}_1)$ are the maximum likelihood estimators of α_z in each treatment group under the prediction model for $\Pr(R_i = 1|X_i; \alpha_z) = w(X_i; \alpha_z)$. Typically, $w(\cdot)$ would be chosen to be a logistic function. Under the assumption (4) and the true model for $w(X_i; \alpha_z)$, there exists a unique solution $\hat{\beta}$, and $U(\beta_{\text{true}}, \hat{\alpha})$ and $(\hat{\beta} - \beta_{\text{true}})$ are asymptotically normal with a mean of zero and respective asymptotic variances C and $(\Gamma^{-1})C\,(\Gamma^{-1})^T$ that can be consistently estimated by $\hat{C}$ and $(\hat{\Gamma}^{-1})\hat{C}(\hat{\Gamma}^{-1})^T$, where $\hat{\Gamma} = \partial U(\hat{\beta}, \hat{\alpha})/\alpha\beta^T$, $\hat{C} = \hat{A} - \hat{B}\hat{\Omega}\hat{B}^T$, $\hat{A} = \Sigma U_i(\hat{\beta}, \hat{\alpha}) U_i(\hat{\beta}, \hat{\alpha})^T$, $\hat{B} = \partial U(\hat{\beta}, \hat{\alpha})/\partial\alpha^T$, and $\hat{\Omega}$ is the estimate of Ω based on the observed information from the likelihood for the prediction model $w(X_i; \alpha_z)$. The above variance estimator for β can be expressed as $(\hat{\Gamma}^{-1})\hat{A}(\Gamma^{-1})^T(\hat{\Gamma}^{-1})\hat{B}\hat{\Omega}\hat{B}^T(\hat{\Gamma}^{-1})^T$. The first term $(\hat{\Gamma}^{-1})\hat{A}(\hat{\Gamma}^{-1})^T$ is the usual robust (sandwich) variance estimate without accounting for the variability of weights, thus, the estimator $\hat{\beta}$ that uses the estimated weights is at least as efficient as the one based on the usual robust variance estimate.

Clinical Study

Oncology

Oncology 2005;68:171–178
DOI: 10.1159/000086771

Received: February 4, 2004
Accepted after revision: June 24, 2004
Published online: July 7, 2005

An Early Phase II Study of S-1 in Patients with Metastatic Pancreatic Cancer

Hideki Ueno Takuji Okusaka Masafumi Ikeda Yoriko Takezako
Chigusa Morizane

Hepatobiliary and Pancreatic Oncology Division, National Cancer Center Hospital, Tokyo, Japan

Key Words
Chemotherapy · Pancreatic cancer · Phase II study · S-1 · Pharmacokinetics

Abstract
Objective: The aim of this study was to evaluate the efficacy and toxicity of S-1 in patients with metastatic pancreatic cancer. ***Methods:*** Patients were required to have a histological diagnosis of pancreatic adenocarcinoma with measurable metastatic lesions, and no prior chemotherapy. S-1 was administered orally at 40 mg/m^2 twice daily for 28 days with a rest period of 14 days as one course. Administration was repeated until the appearance of disease progression or unacceptable toxicity. A pharmacokinetic study was done on day 1 in the initial 8 patients. ***Results:*** Nineteen patients were entered into this study. Four patients (21.1%) achieved a partial response with a 95% confidence interval of 6.1–45.6%. No change was noted in 10 patients (52.6%), and progressive disease in 5 patients (26.3%). The median survival time was 5.6 months with a one-year survival rate of 15.8%. The major adverse events were gastrointestinal toxicities such as nausea and anorexia, though most of them were tolerable and reversible. There were no large differences in the pharmacokinetic parameters of S-1 in patients with pancreatic cancer and those in patients with other cancers. ***Conclusion:*** S-1 is active and tolerated in patients with metastatic pancreatic cancer, which will be confirmed in the following large-scale phase II study.

Introduction

Pancreatic cancer is among the most lethal of all solid tumors. More than 80% of patients have unresectable disease at diagnosis, and even if resection is performed, the recurrence rate is extremely high. Consequently, only ≤ 5% of all patients with pancreatic cancer survive 5 years after diagnosis [1]. Although pancreatic cancer has been considered as a chemotherapy-resistant tumor, recent studies have demonstrated that gemcitabine is an effective tool for the palliation of symptoms and prolonging survival in patients with advanced pancreatic cancer. However, single-agent gemcitabine has provided limited benefit, with objective response rates of less than 15% and a median survival of less than 6 months [2–8]. Therefore, to improve the prognosis of patients with pancreatic cancer, there is a clear need to identify a new effective chemotherapeutic regimen.

KARGER
Fax + 41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

0030–2414/05/0683–0171$22.00/0

Accessible online at:
www.karger.com/ocl

Hideki Ueno, MD
Hepatobiliary and Pancreatic Oncology Division, National Cancer Center Hospital
5-1-1 Tsukiji, Chuo-ku
Tokyo 104-0045 (Japan)
Tel. +81 3 3542 2511, Fax +81 3 3542 3815, E-Mail hueno@ncc.go.jp

S-1 is an oral anticancer drug, which consists of tegafur (FT) as a prodrug of 5-fluorouracil (5-FU), 5-chloro-2,4-dihydroxypyridine (CDHP) and potassium oxonate (Oxo) [9]. The drug has been developed to improve the tumor-selective toxicity of 5-FU by two biochemical modulators, CDHP and Oxo. CDHP is a competitive inhibitor of dihydropyrimidine dehydrogenase involved in the degradation of 5-FU, and maintains efficacious 5-FU concentrations in plasma and tumor tissues [10]. Oxo, a competitive inhibitor of orotate phosphoribosyltransferase, inhibits phosphorylation of 5-FU in the gastrointestinal tract, and reduces the serious gastrointestinal toxicity associated with 5-FU [11]. S-1 has already demonstrated a potent antitumor effect in clinical studies on various solid tumors [12–18]. The response rates in the late phase II studies for advanced colorectal cancer, non-small cell lung cancer, and head and neck cancer were 35, 22, and 29%, respectively [12–14]. In particular, an excellent antitumor effect was demonstrated in the two late phase II studies for advanced gastric cancer, which resulted in response rates of 49 and 44%, respectively [15, 16]. In these late phase II studies, S-1 was administered at a dose of 80 mg/m^2/day for 28 consecutive days followed by a rest period of 14 days, based on the experience of the early phase II studies [17, 18]. The major adverse events recognized in these studies were myelosuppression and gastrointestinal toxicities, though most of them were tolerable and reversible. According to these findings, the commercial availability of S-1 for the treatment of patients with gastric cancer, colorectal cancer and head and neck cancer has been approved in Japan.

As for pancreatic cancer, although the preclinical antitumor efficacy of S-1 on human pancreatic cancer xenografts implanted into nude rats has been reported [19], its clinical activity against pancreatic cancer has not been evaluated. As it is available in an oral form, S-1 has a potential advantage as far as the convenience of the patients is concerned, especially in terms of quality of life. This is very important in pancreatic cancer patients, because the remaining life span of these patients is generally short. Thus, we conducted an early phase II study to evaluate the antitumor effect and safety of S-1 in patients with metastatic pancreatic cancer.

Patients and Methods

Study Patients

All patients were required to show histologically proven pancreatic adenocarcinoma with measurable metastatic lesions. Additional criteria included the following: no history of prior antitumor treatment except pancreatic resection; 20–74 years of age; Karnofsky performance status of 80–100 points; estimated life expectancy ≥2 months; adequate marrow function (white blood cell count 4,000–12,000/mm^3, platelet count ≥100,000/mm^3, hemoglobin level ≥10.0 g/dl), adequate renal function (normal serum creatinine level), adequate liver function (total bilirubin level ≤3 times upper normal limit, transaminases levels ≤2.5 times upper normal limit), and written informed consent from the patients. Patients were excluded if there was a history of drug hypersensitivity, serious complications, symptoms attributable to brain metastasis, active secondary cancer, active infection, marked pleural or peritoneal effusion, and watery diarrhea. Pregnant or lactating women were also excluded. The study was performed in accordance with the Declaration of Helsinki, approved by the institutional review board at the National Cancer Center Hospital, and conducted in accordance with the Good Clinical Practice guidelines in Japan.

Treatment Schedule

S-1 was administered orally at 40 mg/m^2 twice daily after breakfast and dinner. Three initial doses were established according to the body surface area (BSA) as follows: BSA <1.25 m^2, 80 mg/day; 1.25 m^2 ≤ BSA <1.50 m^2, 100 mg/day; and 1.50 m^2 ≤ BSA, 120 mg/day. S-1 was administered at the respective dose for 28 days, followed by a 14-day rest period. This schedule was repeated every 6 weeks until the occurrence of disease progression, unacceptable toxicities, or the patient's refusal to continue. If grade 3 or higher hematological toxicity or grade 2 or higher nonhematological toxicity was observed, the temporary interruption of S-1 and/or the dose reduction by 20 mg/day was allowed (minimum dose, 80 mg/day). Unless adverse events appeared, to enhance the pharmacological effect, the rest period was shortened to 7 days or the dose was gradually escalated in the next course (maximum dose, 150 mg/day), or both were permitted according to the judgment of individual physicians. If a rest period of more than 28 days was required, the patient was withdrawn from the study. During the treatment, patients maintained a daily journal to record their S-1 intake and any adverse events experienced. S-1 was provided by Taiho Pharmaceutical Co. Ltd. (Tokyo, Japan).

Evaluation of Response and Safety

The response was assessed using computed tomography scan or magnetic resonance imaging in each course according to the Japan Society for Cancer Therapy Criteria [20], which is basically similar to the World Health Organization Criteria. Briefly, complete response was defined as the complete disappearance of all measurable and assessable lesions for at least 4 weeks. Partial response was defined as a ≥50% reduction in the sum of the products of the greatest perpendicular diameters of all measurable lesions for at least 4 weeks. No change was defined as a <50% reduction or a <25% increase in the products of the greatest perpendicular diameters of all lesions for at least 4 weeks. Progressive disease (PD) was defined as a ≥25% increase or the appearance of new lesions. Primary pancreatic lesions were considered to be assessable but not measurable lesions, because it is difficult to measure the size of primary pancreatic lesions accurately [21].

Physical examination, complete blood cell counts, biochemistry tests, and urinalysis were assessed weekly during the treatment. Adverse events were evaluated according to the National Cancer Institute Common Toxicity Criteria version 2.0. An external review committee confirmed the objective responses and adverse events.

Table 1. Patient characteristics (n = 19)

Characteristics		Patients	%
Gender			
Male		13	68
Female		6	32
Median age, years (range)	61 (45–73)		
Karnofsky performance status			
100 points		2	11
90 points		16	84
80 points		1	5
Median first dose, mg/m^2 (range)	36.7 (33.7–39.9)		
History of pancreatectomy		1	5
Sites of metastasis			
Liver		15	79
Distant lymph node		3	16
Lung		3	16
Peritoneum		1	5
Median CEA, ng/ml (range)	8.6 (0.4–121)		
Median CA 19-9, U/ml (range)	4,033 (1–155,400)		

Pharmacokinetics

A pharmacokinetic study was performed in the first 8 patients enrolled in the study. Blood (5 ml) was collected with a heparinized syringe on day 1 of the first course before and 1, 2, 4, 6, 8, 10, and 12 h after the administration of S-1. Plasma was separated by centrifugation, and stored at –20 °C until analysis. Plasma concentrations of FT, 5-FU, CDHP, and Oxo were quantified as reported previously [22]. FT was quantified by high-performance liquid chromatography with UV detection, and 5-FU, CDHP, and Oxo were quantified by gas chromatography-negative ion chemical ionization mass spectrometry.

Pharmacokinetic parameters, maximum plasma concentration (C_{max}, ng/ml), time to reach C_{max} (T_{max}, h), area under the concentration versus time curve zero to infinity ($AUC_{0-\infty}$, ng·h/ml), and elimination half-life ($T_{1/2}$, h) were calculated by a noncompartment model in Win-Nonlin Version 3.1 (Pharsight, Apex, NC, USA).

Statistical Analysis

The response duration was calculated from the day of the first demonstration of response until PD; time to progression was calculated from the date of study entry until documented PD; overall survival time was calculated from the date of study entry to the date of death or the date of the last follow-up. Median probability of survival and the median time to progression were estimated by the Kaplan-Meier method. Compliance was calculated for all the courses using the ratio of the total dose actually administered to the scheduled dose. Analysis was planned to be carried out when 19 patients were enrolled. In this study, the threshold rate was defined as 5% and the expected rate was set as 15%. If the lower limit of the 90% confidence interval exceeded the 5% threshold (objective response in 4 or more of the 19 patients), S-1 was judged to be effective and we would proceed to the next large-scale study. If the upper limit of the 90% confidence interval did not exceed the expected rate of 15% (no objective response in the 19 patients), S-1 was judged to be ineffective and the study was to be ended. If response was confirmed in 1–3 of the 19 patients, whether to proceed to the next study or not was judged based on the safety and survival data from the present study.

Results

Patients

Nineteen consecutive patients with metastatic pancreatic cancer were enrolled in this study between June 2000 and January 2001 at the National Cancer Center Hospital. All patients were eligible and assessable for responses and adverse events. The patient characteristics are shown in table 1. The Karnofsky performance status was 80–100 points in all patients, and 18 of the 19 showed a Karnofsky performance status of ≥ 90. Before chemotherapy, morphine was prescribed for 7 patients due to abdominal and/or back pain.

Treatments

A total of 56 courses were administered to the 19 patients with a median of 2 courses per patient (range, 1–12). The initial administered dose of S-1 was 100 mg/day in 8 patients and 120 mg/day in 11 patients. Dose reduction was required in one patient because of grade 3 nausea, vomiting, and anorexia. The compliance rate of the patients taking S-1 during all the courses was as good as 90%.

Response and Survival

Out of the total of 19 evaluable patients, although no complete response was seen, partial response was obtained in 4 patients, resulting in an overall response rate of 21.1% (95% CI, 6.1–45.6%). No change was noted in 10 patients (52.6%), and PD in 5 patients (26.3%).

Table 2. Characteristics of responding patients (n = 4)

Patient No.	Gender	Age	KPS	History of pancreatectomy	Sites of metastasis	Symptomatic benefits	Response duration days	Survival time days
7	Female	65	90	No	Liver	Not assessable	78	463+
17	Female	61	90	No	Liver	No change	205	253
18	Female	68	90	No	Lung	No change	418	452+
19	Male	63	90	Yes	Abdominal lymph node	Improved[a]	213	448+

[a] Morphine consumption was decreased to ≥ 50% from baseline for 27 weeks without any deterioration of the KPS.

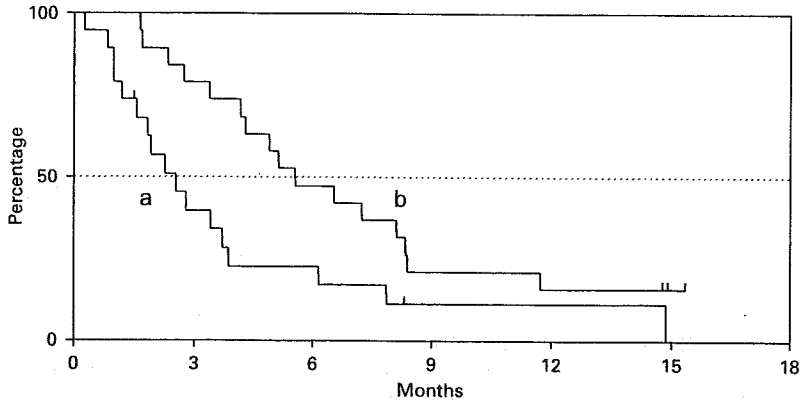

Fig. 1. Time to progression (a), and overall survival time (b).

Responses for each of the target sites were 20.0% (3/15) in liver, 33.3% (1/3) in the distant lymph nodes, and 33.3% (1/3) in lung metastases, respectively. The median time from the date of study entry to the day of the first demonstration of response was 34.5 days (range, 31–35 days) and the median response duration was 7.0 months (range, 2.6–13.9 months). The characteristics of all responders are shown in table 2. The median time to progression was 2.6 months, and the overall median survival was 5.6 months with a one-year survival rate of 15.8% (fig. 1). The serum CA 19-9 level was reduced to less than half in 7 (43.8%) of 16 patients who had a pretreatment level of 100 U/ml or greater.

Safety

S-1 was tolerated in this study. Treatment-related adverse events are listed in table 3. The most common adverse events were nausea (grade ≥ 1, 68.4%) and anorexia (grade ≥ 1, 57.9%), though most of them were tolerable and reversible. Vomiting, stomatitis, diarrhea, and skin rash were generally mild and less frequent, and no serious hepatic or renal toxicities were observed. As to hematological toxicities, grade ≥ 3 neutropenia was noted in only one patient (5.3%), and no grade ≥ 3 thrombocytopenia was observed. Although most patients could be treated as an outpatient without severe adverse events, 3 patients required hospitalization due to grade 3 ileus. Ileus occurred in the first course of treatment in 2 patients, and the remaining one had this event in the sixth course of treatment. However, all of them recovered from ileus after interruption of the S-1 with appropriate treatment. No other severe or unexpected adverse events were noted. Although 2 patients died within 2 months due to rapid disease progression, no treatment-related deaths were observed.

Table 3. Treatment-related adverse events (n = 19): worst grade reported during treatment period

Toxicity	Grade				Grade 1–4	Grade 3–4
	1	2	3	4	%	%
Hematological						
Leukopenia	1	1	0	0	10.5	0
Neutropenia	1	1	1	0	15.8	5.3
Hemoglobin	1	5	1	0	36.8	5.3
Thrombocytopenia	6	0	0	0	31.6	0
Nonhematological						
Nausea	10	0	3	0	68.4	15.8
Vomiting	4	1	1	0	31.6	5.3
Anorexia	6	2	2	1	57.9	15.8
Stomatitis	5	0	0	0	26.3	0
Diarrhea	2	1	1	0	21.1	5.3
Abdominal distension	3	0	2	0	26.3	10.5
Ileus	0	0	3	0	15.8	15.8
Colitis	0	0	2	0	10.5	10.5
Fatigue	3	1	1	0	26.3	5.3
Skin rash	2	1	0	0	15.8	0
Pigmentation	2	2	0	0	21.1	0
Aspartate aminotransferase	3	0	0	0	15.8	0
Alanine aminotransferase	1	2	0	0	15.8	0
Creatinine	0	0	0	0	0	0

Table 4. Pharmacokinetic parameters of FT, 5-FU, CDHP, and Oxo after administration of S-1 (n = 8)

	C_{max} ng/ml	T_{max} h	$AUC_{0-\infty}$ ng·h/ml	$T_{1/2}$ h
FT	1,705±383	2.9±1.2	23,846±9,848	8.9±2.4
5-FU	125.7±46.8	4.0±1.1	680.5±252.1	1.9±0.3
CDHP	217.3±100.6	3.0±1.1	1,139.3±335.7	2.9±0.4
Oxo	48.7±51.1	2.4±1.1	253.3±277.6	2.4±0.8

Parameters are represented as mean ± SD.

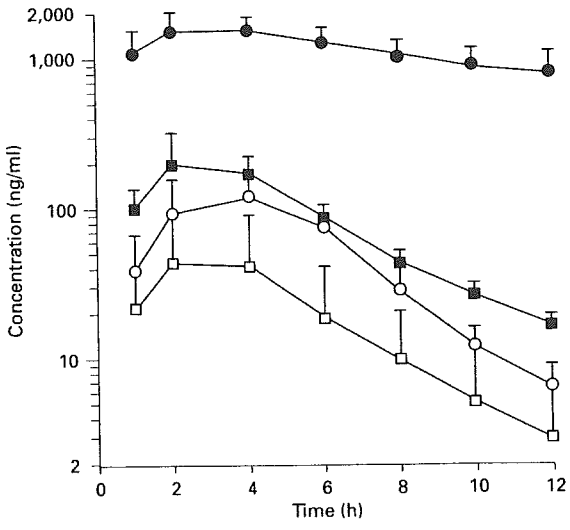

Fig. 2. Plasma concentration-time profiles of FT (●), 5-FU (○), CDHP (■), and Oxo (□) after administration of S-1 (n = 8). The values are expressed as the mean ± SD.

Pharmacokinetics

The pharmacokinetic parameters (C_{max}, T_{max}, $AUC_{0-\infty}$, and $T_{1/2}$) for FT, 5-FU, CDHP, and Oxo are listed in table 4. Plasma concentrations of all compounds peaked between 2 and 4 h after administration. The plasma concentration of FT reached a plateau after C_{max}, which was maintained for 12 h, while 5-FU, CDHP, and Oxo were more rapidly eliminated from the systemic circulation (fig. 2).

Discussion

5-FU, first synthesized 40 years ago, is still one of the most widely used agents for digestive system cancers including pancreatic cancer. Since 5-FU shows a short half-life and a time-dependent effect, its continuous infusion is known to result in a better antitumor effect than bolus injection [23]. A meta-analysis of six randomized trials has demonstrated that the continuous infusion 5-FU is superior to bolus 5-FU with respect to tumor response and survival in metastatic colorectal cancer [24]. As for pancreatic cancer, a recent study by Maisey et al. [25] has reported that the continuous infusion of 5-FU for the treatment of advanced pancreatic cancer results in a response rate of 8.4% and a median survival time of 5.1 months. However, continuous infusion of 5-FU requires a catheter, and is associated with complications, such as infections, and a reduced quality of life. Moreover, patients receiving continuous infusion of 5-FU show disturbance of their circadian rhythms and intraindividual variations in plasma 5-FU levels caused by dihydropyrimidine dehydrogenase, which contribute to limiting the effect of 5-FU. In addition, continuous infusion of 5-FU may cause severe gastrointestinal toxicities such as diarrhea and stomatitis. To overcome these problems, an oral fluoropyrimidine derivative, S-1, was developed on the basis of the biochemical modulation by CDHP, a dihydropyrimidine dehydrogenase inhibitor, and Oxo, a protector against 5-FU-induced gastrointestinal toxicity. Since the antitumor effects of S-1 on various solid cancers have been reported [12–18], we considered that the efficacy of S-1 on pancreatic cancer should also be investigated.

S-1 showed a good objective response rate of 21.1% with a good tumor growth control rate (objective responses plus no change) of 73.7% for metastatic pancreatic cancer patients. In the reported phase II and III studies for pancreatic cancer, single-agent gemcitabine showed response rates ranging from 5.4 to 16.0%, mostly below 15%, and tumor growth control rates ranging from 25.1–72.0%, mostly below 50% [2–8]. Our study also demonstrated a median survival time of 5.6 months with a one-year survival rate of 15.8%, which was comparable to the results of the gemcitabine studies. S-1 was easily administered, and most patients could be treated as outpatients. These results suggest that S-1 has an antitumor effect on metastatic pancreatic cancer.

A pharmacokinetic study of S-1 has already been conducted by Hirata et al [26]. They administered S-1 twice daily at a dose of 80 mg/m^2/day in 12 patients with gastric, colorectal, and breast cancer, and reported that C_{max}, T_{max}, AUC_{0-14}, and $T_{1/2}$ of 5-FU after a single administration of S-1 were 128.5 ± 41.5 ng/ml, 3.5 ± 1.7 h, 723.9 ± 272.7 ng·h/ml, and 1.9 ± 0.4 h, respectively. The pharmacokinetic parameters of 5-FU observed in our study (C_{max}, 125.7 ± 46.8 ng/ml; T_{max}, 4.0 ± 1.1 h; $AUC_{0-\infty}$, 680.5 ± 252.1 ng·h/ml; $T_{1/2}$, 1.9 ± 0.3 h) were similar to those in Hirata's study. The pharmacokinetic parameters of other compounds, FT, CDHP, and Oxo, also did not show a large difference between the two studies. Therefore, our data suggest that there were no large differences between the pharmacokinetic parameters of S-1 in patients with pancreatic cancer and those in patients with other cancers.

Toxicity of S-1 was acceptable in our study. Hematological toxicities were mild, similar to the results of clinical studies of S-1 for other cancers. However, gastrointestinal toxicities such as anorexia and vomiting tended to occur more frequently in our study. Grade ≥ 3 anorexia and vomiting were observed in 4.8 and 1.6% of colorectal cancer patients [12], while grade ≥ 3 anorexia and vomiting were seen in 15.8 and 5.3% of pancreatic cancer patients. Since the pharmacokinetic parameters of S-1 did not differ between subjects with pancreatic cancer and those with other cancers, we speculate that anorexia and vomiting were observed more frequently partly because many patients with pancreatic cancer had disease-related symptoms such as anorexia before treatment. Although phase I studies for S-1 from the Netherlands and the United States described diarrhea as a dose-limiting factor [27, 28], diarrhea was mild and low in incidence in this study, similar to the results of other cancer studies conducted in Japan. However, 3 patients in the current study required hospitalization because of ileus, an observation different from the past Japanese reports. In the United States, an 80-year-old female with gallbladder cancer was reported as developing grade 4 ileus with grade 3 diarrhea after administration of S-1 [28]. In the current study, 1 of the 3 patients had concomitant colitis, while the remaining 2 had no colitis. Although the causes of the ileus were unknown, S-1 may have been the underlying cause, because all patients recovered from ileus after cessation of S-1 with appropriate treatment. Two of the 3 patients had been put on morphine, and showed a tendency towards constipation before the onset, suggesting that the administration of S-1 requires attention to bowel movements.

In this study, since no serious adverse events occurred except the above-described ileus, most patients could be treated as outpatients. The compliance rate of the patients receiving S-1 was as good as 90%. S-1 is an oral anticancer

drug, and has the advantage of being able to treat patients while maintaining their quality of life. Since the prognosis of patients with advanced pancreatic cancer is generally poor, the demonstration in this study of the effectiveness and safety of S-1 (which allows treatment on an outpatient basis) for pancreatic cancer is highly significant. As the toxicity of S-1 is relatively mild, S-1 can be used in combination with other anticancer drugs. Combination therapy with S-1 and cisplatin has already been conducted for gastric cancer, and an excellent response rate of 76% was reported in a phase II study [29], which encourages the expectation of a future combination therapy with S-1 and other anticancer drugs including gemcitabine for advanced pancreatic cancer as well.

In conclusion, although this study had a small patient population, S-1 showed a promising antitumor activity with tolerable toxicity in metastatic pancreatic cancer patients. As an oral medication, S-1 offers a potential advantage as far as patient convenience is concerned, especially in terms of the patients' quality of life. We are currently conducting a multi-institutional late phase II study of S-1 for metastatic pancreatic cancer to confirm the results in this study.

Acknowledgements

We are grateful to Drs. M. Kurihara, S. Matsuno, O. Ishikawa, and T. Taguchi for their kind advice, Drs. H. Saisho, N. Moriyama, and W. Koizumi for extramural review, and T. Tomiazawa for good support. We also thank T. Tahara, T. Tsuruda, A. Fukushima, M. Noguchi, and R. Azuma, PhD, for their assistance in data management, and K. Kuwata for the pharmacokinetic analyses. The study was supported by Taiho Pharmaceutical Co., Ltd.

References

1 Evans DB, Abbruzzese JL, Willett CG: Cancer of the pancreas; in De Vita VT Jr, Hellman S, Rosenberg SA (eds): Principles and Practice of Oncology, ed 6. Philadelphia, Lippincott, 2001, pp 1126–1161.
2 Burris HA 3rd, Moore MJ, Andersen J, Green MR, Rothenberg ML, Modiano MR, Cripps MC, Portenoy RK, Storniolo AM, Tarassoff P, Nelson R, Dorr FA, Stephens CD, Von Hoff DD: Improvements in survival and clinical benefit with gemcitabine as first-line therapy for patients with advanced pancreas cancer: A randomized trial. J Clin Oncol 1997;15:2403–2413.
3 Casper ES, Green MR, Kelsen DP, Heelan RT, Brown TD, Flombaum CD, Trochanowski B, Tarassoff PG: Phase II trial of gemcitabine (2,2′-difluoro-deoxycytidine) in patients with adenocarcinoma of the pancreas. Invest New Drugs 1994;12:29–34.
4 Carmichael J, Fink U, Russell RC, Spittle MF, Harris AL, Spiessi G, Blatter J: Phase II study of gemcitabine in patients with advanced pancreatic cancer. Br J Cancer 1996;73:101–105.
5 Bramhall SR, Rosemurgy A, Brown PD, Bowry C, Buckels JAC: Marimastat as first-line therapy for patients with unresectable pancreatic cancer: A randomized trial. J Clin Oncol 2001; 19:3447–3455.
6 Bramhall SR, Schulz J, Nemunaitis J, Brown PD, Baillet M, Buckels JAC: A double-blind placebo-controlled, randomised study comparing gemcitabine and marimastat with gemcitabine and placebo as first line therapy in patients with advanced pancreatic cancer. Br J Cancer 2002;87:161–167.
7 Berlin JD, Catalano P, Thomas JP, Kugler JW, Haller DG, Benson AB III: Phase III study of gemcitabine in combination with fluorouracil versus gemcitabine alone in patients with advanced pancreatic carcinoma: Eastern Cooperative Oncology Group Trial E2297. J Clin Oncol 2002;20:3270–3275.
8 Colucci G, Giuliani F, Gebbia V, Biglietto M, Rabitti P, Uomo G, Cigolari S, Testa A, Maiello E, Lopez M: Gemcitabine alone or with cisplatin for the treatment of patients with locally advanced and/or metastatic pancreatic carcinoma: A prospective, randomized phase III study of the Gruppo Oncologico dell'Italia Meridionale. Cancer 2002;94:902–910.
9 Shirasaka T, Shimamoto Y, Ohshimo H, Yamaguchi M, Kato T, Yonekura K, Fukushima M: Development of a novel form of an oral 5-fluorouracil derivative (S-1) directed to the potentiation of the tumor selective cytotoxicity of 5-fluorouracil by two biochemical modulators. Anticancer Drugs 1996;7:548–557.
10 Tatsumi K, Fukushima M, Shirasaka T, Fujii S: Inhibitory effects of pyrimidine, barbituric acid and pyridine derivatives on 5-fluorouracil degradation in rat liver extracts. Jpn J Cancer Res 1987;78:748–755.
11 Shirasaka T, Shimamoto Y, Fukushima M: Inhibition by oxonic acid of gastrointestinal toxicity of 5-fluorouracil without loss of its antitumor activity in rats. Cancer Res 1993;53:4004–4009.
12 Ohtsu A, Baba H, Sakata Y, Mitachi Y, Horikoshi N, Sugimachi K, Taguchi T: Phase II study of S-1, a novel oral fluoropyrimidine derivative, in patients with metastatic colorectal carcinoma. S-1 Cooperative Colorectal Carcinoma Study Group. Br J Cancer 2000;83: 141–145.
13 Kawahara M, Furuse K, Segawa Y, Yoshimori K, Matsui K, Kudoh S, Hasegawa K, Niitani H: Phase II study of S-1, a novel oral fluorouracil in advanced non-small-cell lung cancer. Br J Cancer 2001;85:939–943.
14 Inuyama Y, Kida A, Tsukuda M, Kohno N, Satake B: Late phase II study of S-1 in patients with advanced head and neck cancer. Gan To Kagaku Ryoho 2001;28:1381–1390.
15 Sakata Y, Ohtsu A, Horikoshi N, Sugimachi K, Mitachi Y, Taguchi T: Late phase II study of novel oral fluoropyrimidine anticancer drug S-1 (1 M tegafur-0.4 M gimestat-1 M otastat potassium) in advanced gastric cancer patients. Eur J Cancer 1998;34:1715–1720.
16 Koizumi W, Kurihara M, Nakano S, Hasegawa K: Phase II study of S-1, a novel oral derivative of 5-fluorouracil, in advanced gastric cancer. For the S-1 Cooperative Gastric Cancer Study Group. Oncology 2000;58:191–197.
17 Sugimachi K, Maehara Y, Horikoshi N, Shimada Y, Sakata Y, Mitachi Y, Taguchi T: An early phase II study of oral S-1, a newly developed 5-fluorouracil derivative for advanced and recurrent gastrointestinal cancers. The S-1 Gastrointestinal Cancer Study Group. Oncology 1999;57:202–210.
18 Furuse K, Kawahara M, Hasegawa K, Kudoh S, Takada M, Sugiura T, Ichinose Y, Fukuoka M, Ohashi Y, Niitani H: Early phase II study of S-1, a new oral fluoropyrimidine, for advanced non-small-cell lung cancer. Int J Clin Oncol 2001;6:236–241.

19 Fukushima M, Satake H, Uchida J, Shimamoto Y, Kato T, Takechi T, Okabe H, Fujioka A, Nakano K, Ohshimo H, Takeda S, Shirasaka T: Preclinical antitumor efficacy of S-1: A new oral formulation of 5-fluorouracil on human tumor xenografts. Int J Oncol 1998;13:693–698.
20 Japan Society for Cancer Therapy: Criteria for the evaluation of the clinical effects of solid cancer chemotherapy. J Jpn Soc Cancer Ther 1993;28:101–130.
21 Aoki K, Okada S, Moriyama N, Ishii H, Nose H, Yoshimori M, Kosuge T, Ozaki H, Wakao F, Mukai K: Accuracy of computed tomography in determining pancreatic cancer tumor size. Jpn J Clin Oncol 1994;24:85–87.
22 Matsushima E, Yoshida K, Kitamura R, Yoshida K: Determination of S-1 (combined drug of tegafur, 5-chloro-2,4-dihydroxypyridine and potassium oxonate) and 5-fluorouracil in human plasma and urine using high-performance liquid chromatography and gas chromatography-negative ion chemical ionization mass spectrometry. J Chromatogr B 1997;691:95–104.
23 Lokich JJ, Ahlgren JD, Gullo JJ, Philips JA, Fryer JG: A prospective randomized comparison of continuous infusion fluorouracil with a conventional bolus schedule in metastatic colorectal carcinoma: A Mid-Atlantic Oncology Program Study. J Clin Oncol 1989;7:425–432.
24 Meta-analysis Group In Cancer: Efficacy of intravenous continuous infusion of fluorouracil compared with bolus administration in advanced colorectal cancer. J Clin Oncol 1998; 16:301–308.
25 Maisey N, Chau I, Cunningham D, Norman A, Seymour M, Hickish T, Iveson T, O'Brien M, Tebbutt N, Harrington A, Hill M: Multicenter randomized phase III trial comparing protracted venous infusion (PVI) fluorouracil (5-FU) with PVI 5-FU plus mitomycin in inoperable pancreatic cancer. J Clin Oncol 2002;20: 3130–3136.
26 Hirata K, Horikoshi N, Aiba K, Okazaki M, Denno R, Sasaki K, Nakano Y, Ishizuka H, Yamada Y, Uno S, Taguchi T, Shirasaka T: Pharmacokinetic study of S-1, a novel oral fluorouracil antitumor drug. Clin Cancer Res 1999;5:2000–2005.
27 van Groeningen CJ, Peters GJ, Schornagel JH, Gall H, Noordhuis P, de Vries MJ, Turner SL, Swart MS, Pinedo HM, Hanauske AR, Giaccone G: Phase I clinical and pharmacokinetic study of oral S-1 in patients with advanced solid tumors. J Clin Oncol 2000;18:2772–2779.
28 Cohen SJ, Leichman CG, Yeslow G, Beard M, Proefrock A, Roedig B, Damle B, Letrent SP, DeCillis AP, Meropol NJ: Phase I and pharmacokinetic study of once daily oral administration of S-1 in patients with advanced cancer. Clin Cancer Res 2002;8:2116–2122.
29 Ohtsu A, Boku N, Nagashima F, Koizumi W, Tanabe S, Saigenji K, Muro K, Matsumura K, Shirao K: A phase I/II study of S-1 plus cisplatin (CDDP) in patients (pts) with advanced gastric cancer (AGC). Proc Am Soc Clin Oncol 2001;20:656.

Clinical Study

Oncology

Oncology 2005;69:421–427
DOI: 10.1159/000089997

Received: February 8, 2005
Accepted after revision: August 28, 2005
Published online: November 25, 2005

A Phase I Study of Combination Chemotherapy with Gemcitabine and Oral S-1 for Advanced Pancreatic Cancer

Hideki Ueno[a] Takuji Okusaka[a] Masafumi Ikeda[a] Yoriko Ishiguro[a]
Chigusa Morizane[a] Junichi Matsubara[a] Junji Furuse[b] Hiroshi Ishii[b]
Michitaka Nagase[b] Kohei Nakachi[b]

[a]Hepatobiliary and Pancreatic Oncology Division, National Cancer Center Hospital, Tokyo, and
[b]Hepatobiliary and Pancreatic Oncology Division, National Cancer Center Hospital East, Kashiwa, Japan

Key Words
Pancreatic cancer · 5-Fluorouracil · Gemcitabine · S-1

Abstract
Objective: The aim of this study was to determine the maximum-tolerated dose and dose-limiting toxicity (DLT) of combination therapy with gemcitabine and S-1 in patients with advanced pancreatic cancer. ***Methods:*** Chemotherapy-naive patients with histologically or cytologically proven unresectable or metastatic pancreatic cancer were enrolled. The patients received gemcitabine intravenously over 30 min on days 1 and 8 and S-1 orally twice daily from days 1 to 14. Cycles were repeated every 21 days until disease progression. Patients were scheduled to receive gemcitabine (mg/m^2/week) and S-1 (mg/m^2/day) at four dose levels: 800/60 (level 1), 1,000/60 (level 2), 1,000/70 (level 3) and 1,000/80 (level 4). ***Results:*** Eighteen patients were enrolled in this study. The maximum-tolerated dose was not reached even at the highest dose level (level 4) because only 2 of the 6 patients at this level experienced DLT. The DLTs were neutropenia and rash. Six (33%) of the 18 patients achieved a partial response and median overall survival time was 7.6 months. ***Conclusions:*** Combination chemotherapy with gemcitabine and S-1 was well tolerated and showed good antitumor activity in the treatment of pancreatic cancer. We recommend a gemcitabine dose of 1,000 mg/m^2/week and an S-1 dose of 80 mg/m^2/day in further studies with this schedule.

Introduction

Pancreatic cancer is a fatal disease, with a 5-year survival rate of less than 5% [1]. Surgery remains the only curative option for patients with this disease, but the vast majority of patients unfortunately present with advanced, unresectable tumors. Effective non-surgical treatment is therefore needed to improve the outcome in patients with pancreatic cancer.

A randomized controlled study demonstrated that gemcitabine, a nucleoside analogue, is effective in palliating symptoms and prolonging survival in patients with advanced pancreatic cancer: gemcitabine showed a statistically significant advantage both in clinical benefit response (23.8 vs. 4.8%, p = 0.0022) and in median survival (5.65 vs. 4.41 months, p = 0.0025) compared with weekly bolus 5-fluorouracil (5-FU) [2]. Single-agent gemcitabine is currently accepted worldwide as first-line therapy for advanced pancreatic cancer. Nevertheless, there is substantial room for improvement in chemotherapy for pancreatic cancer, because single-agent gemcitabine pro-

KARGER
Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

0030–2414/05/0695–0421$22.00/0

Accessible online at:
www.karger.com/ocl

Hideki Ueno, MD
Hepatobiliary and Pancreatic Oncology Division, National Cancer Center Hospital
5-1-1 Tsukiji, Chuo-ku
Tokyo, 104-0045 (Japan)
Tel. +81 3 3542 2511, Fax +81 3 3542 3815, E-Mail hiueno@ncc.go.jp

vides only limited benefit, with objective response rates of less than 15% and a median survival of less than 6 months [2–5].

S-1 is an oral fluoropyrimidine derivative that combines tegafur with two modulators of 5-FU metabolism, 5-chloro-2,4-dihydroxypyridine and potassium oxonate [6]. 5-Chloro-2,4-dihydroxypyridine is a competitive inhibitor of dihydropyrimidine dehydrogenase, which is involved in the degradation of 5-FU, and acts to maintain efficacious concentrations of 5-FU in plasma and tumor tissues [7]. Potassium oxonate, a competitive inhibitor of orotate phosphoribosyltransferase, inhibits the phosphorylation of 5-FU in the gastrointestinal tract, reducing the serious gastrointestinal toxicity associated with 5-FU [8]. The efficacy of S-1 has already been demonstrated in a variety of solid tumors: the response rates for advanced gastric cancer, colorectal cancer and non-small cell lung cancer in the phase II studies conducted in Japan were 49, 35 and 22%, respectively [9–11]. Recently, the clinical efficacy of S-1 against pancreatic cancer has also been investigated. We conducted an early phase II study of S-1 for metastatic pancreatic cancer and reported that 4 (21.1%) of 19 patients achieved a partial response, with mild toxicity [12]. Hayashi et al. [13] performed a pilot study of single-agent S-1 or S-1 plus cisplatin combination therapy in patients with advanced pancreatic cancer and reported that 3 (20.0%) of the 15 patients or 8 (57.1%) of the 14 patients showed a partial response.

Since S-1 shows a favorable toxicity profile and activity in various solid tumors, including pancreatic cancer, we decided to investigate whether combination therapy with gemcitabine and S-1 is an effective chemotherapeutic regimen for pancreatic cancer. Although many clinical studies of gemcitabine in combination with fluoropyrimidines such as 5-FU, uracil/tegafur and capecitabine have been reported [14–22], little information is available on the combination of gemcitabine and S-1. Thus, we conducted a phase I study to determine the maximum-tolerated dose (MTD) and dose-limiting toxicity (DLT) of gemcitabine and S-1 combination therapy in patients with unresectable or metastatic pancreatic cancer.

Patients and Methods

Patient Selection

Patients were considered eligible if they met the following criteria: histologically or cytologically proven pancreatic adenocarcinoma, unresectable locally advanced or metastatic disease, naive to chemotherapy, Eastern Cooperative Oncology Group performance status of 0–2, age between 20 and 74 years, life expectancy of ≥8 weeks, and adequate organ function defined as white blood cell count ≥4,000/mm^3, neutrophil count ≥2,000/mm^3, platelet count ≥100,000/mm^3, hemoglobin ≥9.0 g/dl, serum creatinine ≤the upper limit of normal, serum albumin ≥3.0 g/dl, total bilirubin ≤2.0 mg/dl, and aspartate aminotansferase and alanine aminotransferase levels ≤2.5 times the upper limit of normal or ≤5 times the upper limit of normal if liver metastases or biliary drainage were present. The exclusion criteria were severe complications, such as infection, heart disease and renal disease (in this study we did not define in detail the exclusion criteria in relation to severe complications), metastasis to the central nervous system, marked pleural effusion or ascites, and watery diarrhea. Pregnant or lactating women were also excluded. Written informed consent was obtained from all patients. This study was approved by the institutional review board at the National Cancer Center and conducted in accordance with the Declaration of Helsinki.

Treatment Plan

This was an open-label, two-center, single-arm phase I study. Gemcitabine (Eli Lilly Japan K.K., Kobe, Japan) was administered as a 30-min intravenous infusion weekly for 2 weeks followed by a 1-week rest. S-1 (Taiho Pharmaceutical Co., Ltd., Tokyo, Japan) was administered orally twice daily from day 1 to day 14 followed by a 1-week rest. The treatment cycles were repeated every 3 weeks until disease progression or unacceptable toxicity occurred. If patients experienced leucopenia <2,000/mm^3, neutropenia <1,000/mm^3, thrombocytopenia <70,000/mm^3, total bilirubin >2.0 mg/dl or aspartate aminotansferase and alanine aminotransferase levels >5 times the upper limit of normal, both gemcitabine and S-1 were withheld until recovery. If patients experienced DLT, the dose of gemcitabine was reduced by 200 mg/m^2/week and the dose of S-1 was reduced by 10 mg/m^2/day in the subsequent cycle. If a rest period of more than 3 weeks was required because of toxicity, the patient was withdrawn from the study.

Patients were scheduled to receive gemcitabine and S-1 at four dose levels (table 1). At the first dose level (level 1), gemcitabine was administered at a dose of 800 mg/m^2/week and S-1 was administered at 60 mg/m^2/day. At the next dose level (level 2), gemcitabine was increased to 1,000 mg/m^2/week with S-1 kept at the same dose. At each of dose levels 3 and 4, S-1 was increased by 10 mg/m^2/day with gemcitabine kept at 1,000 mg/m^2/week. At least 3 patients were enrolled at each dose level. If DLT was observed in the initial 3 patients, a maximum of 3 additional patients was entered into the same dose level. The MTD was defined as the highest dose level that did not cause DLT in 3 of the 3 or ≥3 of the 6 patients treated at that level during the first two cycles of treatment. DLT was defined as grade 4 leucopenia or neutropenia, febrile neutropenia, grade 4 thrombocytopenia, grade 3 thrombocytopenia requiring transfusion, ≥grade 3 non-hematological toxicity excluding nausea, vomiting, anorexia and fatigue, or any toxicity that necessitated a treatment delay of more than 3 weeks. Toxicity was graded according to the National Cancer Institute Common Toxicity Criteria version 2.0.

Patient Evaluation

Physical examinations, complete blood cell counts, biochemistry tests and urinalyses were performed at least once weekly. Tumor assessment with computed tomographic scan or magnetic resonance imaging and measuring of tumor marker CA 19-9 was performed every two cycles, and tumor response was evaluated by the

Table 1. Dose escalation scheme and DLT

Dose level	Gemcitabine mg/m^2/week	S-1 mg/m^2/day	Patients	DLT events	DLT
1	800	60	3	0	
2	1,000	60	3	0	
3	1,000	70	6	1	grade 4 neutropenia
4	1,000	80	6	2	grade 4 neutropenia grade 3 rash and grade 4 neutropenia

criteria of the Japan Society for Cancer Therapy [23], which are similar to those of the World Health Organization. Briefly, a complete response was defined as the disappearance of all clinical evidence of the tumor for a minimum of 4 weeks. A partial response was defined as a 50% or greater reduction in the sum of the products of two perpendicular diameters of all measurable lesions for 4 weeks or longer without any evidence of new lesions. No change was defined as a reduction of less than 50% or a less than 25% increase in the sum of the products of two perpendicular diameters of all lesions for a minimum of 4 weeks. Progressive disease was defined as an increase of 25% or more in the sum of the products of two perpendicular diameters of all lesions, the appearance of any new lesion, or deterioration in clinical status that was consistent with disease progression. The response duration was calculated from the day of the first sign of a response until disease progression; progression-free survival was calculated from the date of the initiation of treatment until documented disease progression or death due to any cause (whichever occurred first); overall survival time was calculated from the date of treatment initiation to the date of death or the last follow-up. The median probabilities of the progression-free or overall survival periods were estimated by the Kaplan-Meier method.

Results

Patient Characteristics

Between September 2003 and July 2004, 18 patients were enrolled in this study. All of them received at least two cycles of chemotherapy and were evaluable for toxicity and response. Patient characteristics are listed in table 2. All patients had good performance status (0 and 1). Two patients had locally advanced unresectable disease and the remaining 16 had metastatic disease. Before the start of the study, 1 patient had received surgical resection and 3 had undergone biliary drainage for obstructive jaundice. Twelve patients had abdominal and/or back pain at study entry. A total of 125 cycles of chemotherapy was administered, with a median of 6 treatment cycles per patient (range 2–22). It was possible to treat all patients as outpatients after one or two cycles of observation in hospital.

Table 2. Patient characteristics

Characteristics		Patients
Patients enrolled		18
Sex		
Male		13
Female		5
Age, years		
Median	61	
Range	43–72	
ECOG performance status		
0		10
1		8
Body surface area, m^2		
Median	1.58	
Range	1.46–1.97	
Disease stage		
Locally advanced		2
Metastatic		16
Sites of metastatic disease		
Liver		13
Lung		2
Distant lymph nodes		5
Pleura		1

ECOG = Eastern Cooperative Oncology Group.

DLT and Recommended Dose

No DLT was observed at dose levels 1 or 2 (table 1). At dose level 3, 1 patient developed grade 4 neutropenia, which was considered DLT, but the remaining 5 did not develop DLT. At dose level 4, the highest dose level, 2 of the 6 patients exhibited DLTs: 1 had grade 4 neutropenia and the other had grade 3 rash concomitant with grade 4 neutropenia. All DLTs occurred in the first cycle of treatment. The MTD was not reached because only 2 of the 6 patients experienced DLT at dose level 4. Therefore, dose level 4 (gemcitabine dose of 1,000 mg/m^2/week and S-1