Surg Today (2014) 44:2305–2313
DOI 10.1007/s00595-014-0859-7

ORIGINAL ARTICLE

Haloperidol prophylaxis does not prevent postoperative delirium in elderly patients: a randomized, open-label prospective trial

Shinji Fukata · Yasuji Kawabata · Ken Fujisiro · Yuichi Katagawa · Kojiro Kuroiwa · Hirotoshi Akiyama · Yasuhito Terabe · Masahiko Ando · Takashi Kawamura · Hideyuki Hattori

Received: 24 September 2013 / Accepted: 16 December 2013 / Published online: 16 February 2014

Abstract
Purpose Postoperative delirium is the most common postoperative complication in the elderly. The purpose of this study was to evaluate the safety and effectiveness of the preventive administration of low-dose haloperidol on the development of postoperative delirium after abdominal or orthopedic surgery in elderly patients.
Subjects A total of 119 patients aged 75 years or older who underwent elective surgery for digestive or orthopedic disease were included in this study.
Methods Patients were divided into those who did (intervention group, $n = 59$) and did not (control group, $n = 60$) receive 2.5 mg of haloperidol at 18:00 daily for 3 days after surgery; a randomized, open-label prospective study was performed on these groups. The primary end-point was the incidence of postoperative delirium during the first 7 days after the operation.
Results The incidence of postoperative delirium in all patients was 37.8 %. No side effects involving haloperidol were noted; however, the incidences of postoperative delirium were 42.4 and 33.3 % in the intervention and control groups, respectively, which were not significantly different ($p = 0.309$). No significant effect of the treatment was observed on the severity or persistence of postoperative delirium.
Conclusions The preventive administration of low-dose haloperidol did not induce any adverse events, but also did not significantly decrease the incidence or severity of postoperative delirium or shorten its persistence.

Keywords Haloperidol prophylaxis · Postoperative delirium · Elderly patients · Randomized open-label prospective trial · NEECHAM

Introduction

The incidence of diseases in the elderly requiring surgery, such as femoral neck fractures and colorectal cancer, has

S. Fukata · Y. Kawabata · K. Fujisiro · Y. Katagawa
Department of Surgery, National Center for Geriatrics and Gerontology, Obu, Japan

S. Fukata (✉)
Department of Perioperative Medical Care, National Center for Geriatrics and Gerontology, 35 Gengo, Morioka, Obu, Aichi 474-8511, Japan
e-mail: fukatash@ncgg.go.jp

K. Kuroiwa
Department of Surgery, Tokyo Metropolitan Geriatric Hospital, Tokyo, Japan

H. Akiyama
Department of Gastroenterological Surgery, Yokohama City University Graduate School of Medicine, Yokohama, Japan

Y. Terabe
Department of Restorative Medicine, National Center for Geriatrics and Gerontology, Obu, Japan

M. Ando
Center for Advanced Medicine and Clinical Research, Nagoya University Graduate School of Medicine, Nagoya, Japan

T. Kawamura
Kyoto University Health Service, Kyoto, Japan

H. Hattori
Department of Psychiatry, National Center for Geriatrics and Gerontology, Obu, Japan

increased with the aging of the population in Japan [1, 2]. An advanced age is not considered to be a contraindication for surgery worldwide, and surgery is actively performed on patients in their 90 s [3]. However, postoperative management of the elderly is accompanied by many risks. In particular, postoperative delirium, which is the most common postoperative complication [4], represents a major issue in the postoperative treatment of the elderly. It is characterized by a disturbance in consciousness/attentiveness/understanding/cognition and a disorder in the sleep–wake rhythm. It develops suddenly and is likely to vary over the course of a day, inducing various psychiatric symptoms and abnormal behavior [5]. Since postoperative delirium develops in the postoperative recovery phase, it makes postoperative management difficult, disturbs postoperative care and recovery, and is associated with various risks, such as the disturbance of medical care due to the removal of a drip infusion, electrodes and monitors. It also puts an excess burden on medical care workers due to frequent nurse calls and violence, and due to the trauma and fractures caused by tumbles and falls. Moreover, restlessness frequently occurs at night when the nursing staff is shorthanded, which is a serious issue in ward management, and causes excess labor and load on nurses and the patients' families. Postoperative delirium also costs more because it delays the postoperative management and prolongs the hospital stay [6].

Once postoperative delirium develops, only symptomatic treatment with drugs (haloperidol) is currently available, and the effect is insufficient in many cases. A high dose may be necessary, which can markedly influence many patients' physical condition. Therefore, methods to prevent postoperative delirium, to rapidly detect the signs of delirium, to prevent progression to a more severe state and to reduce the persistence of the condition need to be established.

Our previous studies confirmed that the NEECHAM confusion scale (NEECHAM) is useful and objective for the evaluation of postoperative delirium; the incidence of postoperative delirium in 75-year-old or older surgical patients was 55 % when evaluated using the NEECHAM scale. Age, preoperative cognitive dysfunction (a low Mini-Mental State Examination (MMSE) score) and a low preoperative NEECHAM score, but not the surgical department or anesthesia type, were significantly correlated with the development of postoperative delirium, and the incidence of postoperative delirium was higher than 80 % in patients with MMSE and NEECHAM scores lower than 25 and 27 before surgery, respectively, which indicated that these patients represent a high-risk group [7]. We considered that the evaluation of the MMSE and NEECHAM before surgery and the NEECHAM during the postoperative course can facilitate the prevention and early treatment of postoperative delirium in elderly patients.

Pharmacological prevention of postoperative delirium by prophylaxis with antipsychotic medication has been reported in several studies [8]. Oral administration of risperidone and olanzapine reduced the incidence of postoperative delirium in patients undergoing cardiac surgery and hip or knee surgery, respectively [9–11]. However, these drugs are not applicable for many patients, because they are unable to have oral intake for a period of time following surgery. In this respect, either ondansetron or haloperidol, whose intravenous injection has been shown to be effective for the treatment of postoperative delirium [12], can be used. Since the aim of our present study was to incorporate routine drug prophylaxis for surgical patients, ondansetron, whose use is not covered by the health insurance system in Japan, was considered to be unfavorable, leaving haloperidol as the only approved injectable medication that could be used. The effects of haloperidol prophylaxis have already been examined in some studies [13–16], but its efficacy for decreasing the occurrence of postoperative delirium, especially in the elderly, is controversial [13–18].

On the basis of these considerations and the results of our previous studies, we performed a randomized, open-label prospective study to investigate the efficacy and safety of the daily postoperative administration of low-dose haloperidol on postoperative delirium in 75-year-old or older patients who underwent abdominal or orthopedic surgery.

Subjects and methods

Ethical considerations

This study was performed in conformity to the ethical principles based on the Declaration of Helsinki and the 'Ethical Guidelines for Clinical Studies' (notification of the Ministry of Health, Labour and Welfare).

(1) The patients were only included in the study when informed consent was obtained.
(2) The patients' privacy was respected, the secrecy of the recorded results was strictly kept and no information obtained from the study results was used for objectives other than research. To prepare the patient evaluation tables, the patients' privacy protection was sufficiently considered, and patients were identified using identification codes.

Regarding the prophylactic intervention study for postoperative delirium in elderly patients, approval had already been obtained from the Ethics and Conflict of Interest Committee of the National Center for Geriatrics and Gerontology. Other study cooperative institutions started the collection of patients after approval by their respective ethics committee.

Study design and objectives

This was a randomized, open-label prospective trial, and the objective was to evaluate the effect of low-dose haloperidol (2.5 mg/day, for the first 3 days after surgery) on the development, severity and persistence of postoperative delirium, and to evaluate the safety of its preventive intravenous administration to patients 75-year old or older who underwent abdominal or orthopedic surgery.

Patients

The subjects consisted of 121 75-year-old or older patients who underwent elective abdominal surgery under general anesthesia or elective orthopedic surgery under general/spinal anesthesia and gave consent to participate in this study at one of five cooperative institutions (National Center for Geriatrics and Gerontology, Tokyo Metropolitan Geriatric Hospital, Yokohama City University Graduate School of Medicine, Aichi-Saiseikai Hospital and Shizuoka-Saiseikai Hospital) between January 2007 and December 2012. Their age, gender, disease treated with surgery, cognitive function (Mini-Mental State Examination: MMSE [19, 20]), activities of daily living (ADL; Barthel Index [21]), NEECHAM confusion scale (NEECHAM) [22–24], and the presence or absence of psychoneurological complications, urinary incontinence, excitement/hyperkinesia during previous hospitalization and the use of psychotropic drugs before admission were evaluated prior to surgery. Patients, who underwent emergency surgery, had a preoperative NEECHAM score below 20, and with periodic dosing with newly added or switched antipsychotics, antidepressants, hypnotics or anti-Parkinson agents within 2 weeks prior to surgery were regarded as ineligible. Patients previously treated with haloperidol for delirium after surgery before the initiation of postoperative preventive haloperidol administration were also excluded.

Measurements and procedures

Eligible patients were enrolled through an internet website on the morning of postoperative day 1 after obtaining consent, and were automatically assigned at that time to the intervention or non-intervention group on a computer using the age, gender and department as adjustment factors.

Haloperidol 0.5A (2.5 mg) was dissolved in 100 ml of saline and intravenously administered by drip infusion once daily at 18:00 from postoperative days 1 to 3 to the intervention group. The dosing time-point of 18:00 was selected because delirium is more likely to occur at night, and also to recover and maintain the sleep–wake rhythm. Regarding the administration method and dose of haloperidol, an intravenous injection of 5–10 mg of haloperidol was recommended as the first-line treatment for orally untreatable delirious patients in the Guidelines for the Treatment of Delirium published by the Japanese Society of General Hospital Psychiatry [25, 26]. The low dose was set in consideration of the physical characteristics of the elderly and the prophylactic nature of the intervention. The duration of administration was decided based on the previous findings in which the development of delirium increased after 24 h, and because severe symptoms continued for approximately 3 days [7]. The development and severity of postoperative delirium were evaluated for 8 days, from postoperative days 0 to 7, using the NEECHAM score.

The NEECHAM score includes the results of an evaluation of three categories: the cognitive information processing function, behavior and physiological control, and the most unfavorable condition over each 24-h period was regarded as the condition on that day. The maximum score of 30 points decreases as the severity of postoperative delirium increases. Patients with a NEECHAM score of 27 or higher, 25–27, 20–24 and 19 or lower were considered to be non-problematic, at high risk of delirium, with mild delirium and with moderate to severe delirium, respectively. This scale has high internal consistency and high reliability regardless of differences among raters, and has been correlated with the Diagnosis and Statistical Manual of Mental Disorders 4th Edition (DSM-IV) diagnostic criteria [27].

The non-intervention group did not receive preventive treatment, and delirium was evaluated in the same way as in the intervention group (Fig. 1).

When delirium developed, conventional treatments, such as the administration of an intravenous antipsychotic drug (such as haloperidol), were administered in both groups.

Assessment and outcomes

Patient data were collected through the internet website, and the identification of personal information was prevented by coding. After completing the data collection from 121 patients, exceeding the planned number of cases, study team members not involved in the medical care of the patients evaluated and analyzed all baseline data and the results. The development and severity of postoperative delirium were evaluated using the NEECHAM score. When the NEECHAM score decreased to below 20 after surgery, the patient was regarded as having developed postoperative delirium.

The primary endpoint was a lower incidence of postoperative delirium in the intervention group than in the non-intervened control group. The secondary endpoints were the severity and persistence of the postoperative delirium in a time-course analysis of the NEECHAM score during the observation period, and the presence or absence of adverse events assumed to be associated with the intervention.

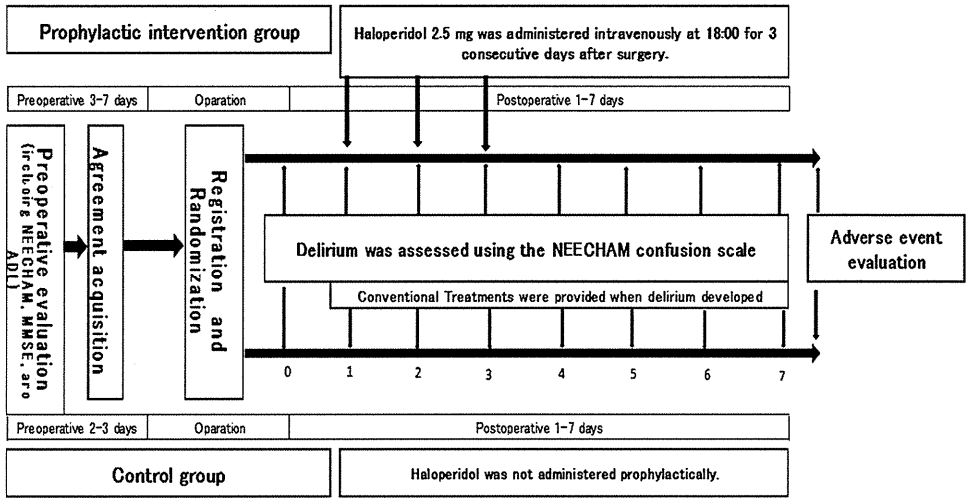

Fig. 1 The study protocol. *NEECHAM* NEECHAM confusion scale, *MMSE* Mini-Mental State Examination, *ADL* Activities of daily living

Discontinuation criteria were withdrawal of consent or a change/discontinuation of treatment requested by the patient or her/his legal representative, the development of National Cancer Institute-Common Toxicity Criteria (NCI-CTC) grade 2 or more severe adverse events associated with haloperidol, difficulty continuing due to severe physical postoperative complications, and judgment that continuing the trial would be difficult by the physician in charge.

Statistical analysis

Eligible patients were randomly assigned 1:1 to the intervention or non-intervention group by the minimization method according to age (<75/≥75 years), gender (female/male), MMSE score (<25/≥25) and institution. The proportions of patients with severe postoperative delirium, defined as at least one episode of a NEECHAM score <20, were compared between the treatment groups using the Chi-square test. This study was designed to have 80 % power to detect a 25 % difference in the proportion of severe postoperative delirium at a two-sided significance level of 0.05. A multivariate logistic regression including the patient age, gender, MMSE score and the preoperative NEECHAM score as covariates was performed to evaluate the effects of the prophylactic haloperidol treatment after adjustment for potential confounding factors. An odds ratio <1 indicated that the factor was protective against severe postoperative delirium. A supportive analysis using the generalized estimating equation regression model was conducted to compare the incidence of severe delirium between treatment groups during the first 7 days after the operation. The statistical analyses were performed using the SAS software program, version 9.3 (SAS Institute, Cary, NC, USA).

Results

In total, 59 and 62 patients were allocated to the prophylactic intervention and control groups, respectively (121 patients overall). The ages of the intervention and control groups were 80.5 ± 0.5 (mean ± standard deviation) and 80.2 ± 0.5 years, respectively. There were 64 male patients (32 each in the intervention and control groups) and 57 female patients (27 and 30, respectively). Abdominal surgery was performed in 107 patients (52 and 55, respectively); orthopedic surgery in nine patients (five and four) and other surgeries (including vascular surgery) were performed in five patients (two and three patients, respectively, in the intervention and control groups). The preoperative MMSE scores in the intervention and control groups were 23.3 ± 0.7 and 23.0 ± 0.7, and the preoperative Barthel Indices were 85.5 ± 3.1 and 84.0 ± 3.0, respectively, with no significant differences observed between the two groups. The preoperative NEECHAM scores were 27.3 ± 0.4 and 28.1 ± 0.4, respectively, with the intervention group having a slightly lower score, although not significant. No significant differences were noted between the two groups in the presence or absence of urinary incontinence, a past medical history of excitement/hyperkinesia or the preoperative use of oral psychotropic drugs, antidepressants, hypnotics or anti-Parkinson agents. The preoperative baseline data of all patients are shown in Tables 1 and 2.

Postoperative NEECHAM measurements were completed in 119 patients (59 and 60 in the intervention and control groups, respectively) because haloperidol was administered to treat delirium on the day of surgery in two of the 62 patients in the control group (Fig. 2). Postoperative delirium (NEECHAM score lower than 20) developed in 45 patients (37.8 %); 20 (33.3 %; 95 %CI 21.7–46.7 %) and 25 patients (42.4 %; 95 %CI 29.6–55.9 %) in the

Table 1 The baseline data of the patients

Preoperative demographics and characteristics	Prophylactic intervention group (*n*:59)	Control group (*n*:62)	*p* value
Age, mean ± SE	80.5 ± 0.5	80.2 ± 0.5	0.723
Male/female ratio	32/27	32/30	0.773
Type of operation			0.852
Abdominal	52	55	
Orthopedic	5	4	
Other	2	3	
MMSE, mean ± SE	23.3 ± 0.7	23.0 ± 0.7	0.740
NEECHAM score, mean ± SE	27.3 ± 0.4	28.1 ± 0.4	0.133
ADL (Barthel index), mean ± SE	85.6 ± 3.1	84.0 ± 3.0	0.736
Urinary incontinence, yes/no	9/50	8/54	0.710
History of excitement, yes/no	1/58	3/59	0.334
Use of anti-Parkinson agents, yes/no	0/59	1/61	0.327
Use of antipsychotic agents, yes/no	1/58	3/59	0.334
Use of antidepressants, yes/no	4/55	1/61	0.154
Use of hypnotics, yes/no	10/49	7/55	0.371

NEECHAM NEECHAM confusion scale, *MMSE* Mini-Mental State Examination, *ADL* Activities of daily living, *Barthel Index* Representative index of the ADL

Table 2 The underlying diseases and type of surgery

Diseases, surgery	Prophylactic intervention group (*n*:59)	Control group (*n*:62)
Abdominal	52	55
Malignancy	36	39
Gastric, gastrectomy/others	11/0	14/2
Colonic, colectomy/others	14/1	14/0
Rectal, LAR/APR/others	2/2/1	4/2/1
Hepatobiliary, hepatectomy/PD/others	1/1/2	1/0/1
Others	1	0
Benign	16	16
Cholelithiasis, cholecystectomy/choledochotomy	0/4	2/2
Abdominal aortic aneurysm, graft	6	7
Others	6	5
Orthopedic	5	4
Others	2	3

LAR low anterior resection, *APR* abdominoperineal resection, Miles operation, *PD* pancreatoduodenectomy

control and intervention groups, respectively. There was no significant effect on the prevention of postoperative delirium ($p = 0.309$).

The postoperative NEECHAM score showed a pattern similar to that previously reported: the score decreased on postoperative day 1 and then gradually increased and returned to the preoperative level on postoperative days 5–7 [7]. The time-course changes in the mean NEECHAM scores in the control and intervention groups are shown in Fig. 3. The mean postoperative NEECHAM scores on postoperative days 1–7 were lower in the intervention group than in the control group, but no significant differences in the severity or incidence of delirium were noted in the intervention group. The mean durations of the persistence of delirium were 1.10 (95 % CI 0.58–1.62 days) and 1.38 days (95 %CI 0.83–1.95 days) in the control and intervention groups, respectively, with no significant difference between them ($p = 0.356$). The incidences of postoperative delirium were 43.2 % (95 %CI 27.1–60.5 %) and 52.8 % (95 %CI 35.5–69.6 %), when the patients were limited to those with a preoperative MMSE <25, and were 64.3 % (95 %CI 35.1–87.2 %) and 66.7 % (95 %CI 43.0–85.4 %), when limited to those with a preoperative NEECHAM score <27, for the control and intervention groups, respectively, which indicated that no significant effect was noted even when patients were limited to those at high risk for postoperative delirium (preoperative MMSE <25 and preoperative NEECHAM <27; $p = 0.415$ and 0.884, respectively).

When a logistic multivariate analysis was performed that included the presence or absence of the intervention as a parameter (Table 3), the incidence of postoperative delirium was significantly higher in patients at an advanced age with low preoperative MMSE and NEECHAM scores [age: odds ratio = 1.12 (for a 1-year increase in age), $p = 0.043$; preoperative MMSE: odds ratio = 1.15 (for a 1-point decrease in the MMSE score), $p = 0.014$; preoperative NEECHAM: odds ratio = 1.23 (for a 1-point decrease in the NEECHAM score), $p = 0.037$]; however, no significant differences associated with gender or the presence or absence of the intervention were noted ($p = 0.953$ and $p = 0.558$, respectively).

Furthermore, when the analysis was conducted after additionally limiting the subjects to those who underwent abdominal surgery, the odds ratio of the prophylactic administration of haloperidol was 1.25 (95 %CI 0.50–3.12), leading to the same conclusion as expected from our previous work which showed no significant difference in the incidence of postoperative delirium between the patients in the department of surgery and orthopedics [7].

To confirm the reliability of the results of the logistic multivariate analysis, an analysis using the generalized estimating equation was performed. The results are shown

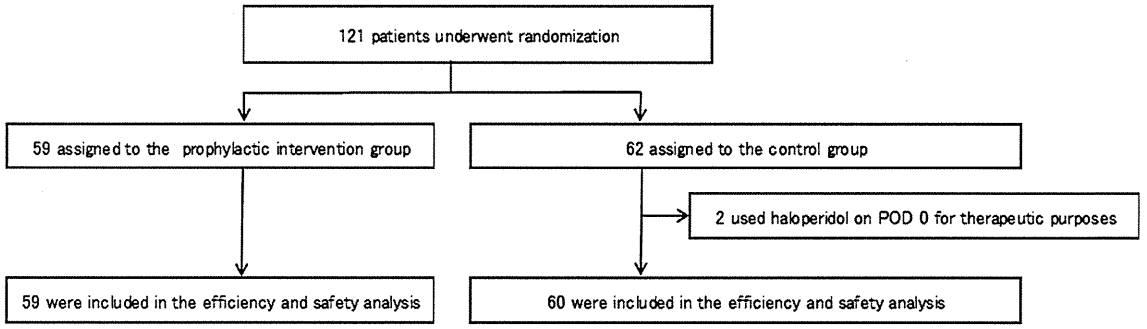

Fig. 2 A flow diagram of the study

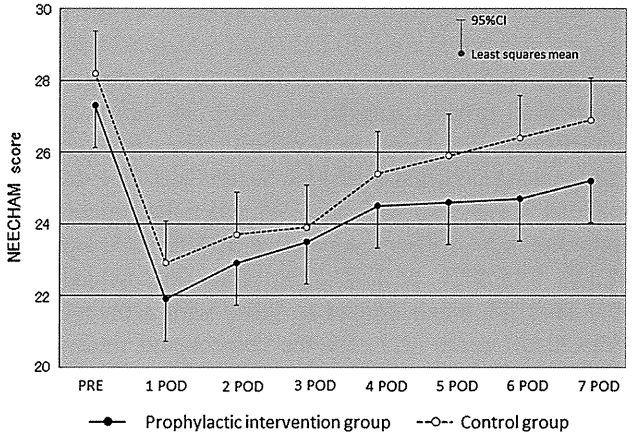

Fig. 3 The preoperative and postoperative changes in the NEECHAM scores in the prophylactic intervention group and control group The preoperative and postoperative (POD1-7) changes in the NEECHAM score in the prophylactic intervention and control groups. The percentages of patients with postoperative delirium (NEECHAM <20) in the prophylactic intervention and the control groups were 42.4 and 33.3 %, respectively (95 % confidence intervals (CI) 29.6–55.9 and 21.7–46.7 %; $p = 0.309$). The duration of the delirium (NEECHAM <20) was 1.38 versus 1.10 days in the intervention and control groups, respectively (95 %CI 0.83–1.95 days and 0.58–1.62 days; $p = 0.356$). *Open circle* shows the least-squares mean score of the control group. *Filled circle* shows the least-squares mean score of the prophylactic intervention group. The *vertical line* in the least-squares mean score shows the 95 %CI

in Table 4, which are consistent with the data shown in Table 3. The odds ratio of the prophylactic administration of haloperidol was 1.48 (95 %CI 0.33–6.71), which indicated no postoperative delirium-decreasing effect ($p = 0.611$). The age, preoperative MMSE score and number of days after surgery were significantly correlated with the development of postoperative delirium.

Regarding postoperative adverse events, grade 2 events were noted in two patients (convulsions and temporary loss of consciousness 2 days after completing the preventive administration of haloperidol and T-tube removal) and grade 3 events occurred in one (no abnormality was noted during haloperidol administration, but the patient fell on

Table 3 The results of the multivariate analysis of the factors affecting the development of postoperative delirium (logistic regression analysis)

	Odds ratio	95 % CI	*p* value
Prophylactic administration of haloperidol	1.30	0.54–3.17	0.558
Age	1.12[a]	1.01–1.27	0.043
Gender (male)	1.03	0.41–2.60	0.953
Preoperative NEECHAM score	1.23[b]	1.01–1.51	0.037
Preoperative MMSE score	1.15[b]	1.03–1.30	0.014

The multivariate logistic regression analysis was performed to include the patient age, gender, preoperative MMSE score and preoperative NEECHAM score as covariates to evaluate the effects of prophylactic haloperidol treatment after adjustments for potential confounding factors. An odds ratio >1 indicated that the factor conferred a greater risk of severe postoperative delirium

NEECHAM NEECHAM confusion scale, *MMSE* Mini-Mental State Examination

[a] For each 1-year increase

[b] For each 1-point decrease in the score

day 7, leading to a femoral neck fracture) in the intervention group. However, the grade 3 event was unlikely to have had any causal relationship with the preventive haloperidol treatment, considering that the half-life of intravenous haloperidol is 14.1 ± 3.2 h [28], and no other adverse events of grade 3 or more were noted.

Discussion

No significant preventive effect of the daily administration of low-dose haloperidol on postoperative delirium was noted in this randomized, open-label prospective study. The incidence of postoperative delirium was not significantly lower in the intervention group than in the control group, and no significant effect was noted on the severity or persistence of the delirium. Moreover, no significant effect was observed even in the group at high

Table 4 The results of the multivariate analysis of the factors affecting the development of postoperative delirium (GEE analysis)

	Odds ratio	95 %CI	*p* value
Prophylactic administration of haloperidol	1.48	0.33–6.71	0.611
Age	1.30[a]	1.07–1.58	0.008
Gender (male)	0.39	0.08–1.86	0.238
Preoperative NEECHAM score	1.05[b]	0.78–1.40	0.752
Preoperative MMSE	1.69[b]	1.05–2.73	0.032
Postoperative day (postoperative days 1–7)	0.67[c]	0.52–0.85	0.001

A supportive analysis using the generalized estimating equation regression model was conducted to compare the incidence of severe delirium between the treatment groups throughout first 7 days after the operation

GEE generalized estimating equation, *NEECHAM* NEECHAM confusion scale, *MMSE* Mini-Mental State Examination

[a] For each 1-year increase

[b] For each 1-point decrease in the score

[c] For each 1-day increase

risk for postoperative delirium with preoperative MMSE and NEECHAM scores below 25 and 27, respectively.

There have been several contradictory reports regarding the efficacy of haloperidol in preventing delirium. Kaneko et al. [13] reported that it significantly decreased the incidence of delirium after surgery involving the digestive organs in a small-scale clinical study, and Wang et al. [14] reported that 12-h continuous preventive administration of haloperidol significantly decreased the incidence of delirium in elderly patients admitted to an ICU, excluding those after cardiac surgery. In contrast, Kalisvaart et al. [15] reported that haloperidol reduced the severity and persistence of delirium, but did not decrease its incidence in patients following orthopedic surgery in a RCT where they received oral preventive treatment with low-dose haloperidol.

The absence of a preventive effect of haloperidol on postoperative delirium may have been due to the low dose used and short administration period. However, although the dose was lower than 5 mg/day for 5 days, as reported by Kaneko et al., the preventive administration of the dose of 2.5 mg/day for 3 days was still markedly higher than the intravenous bolus injection of 0.5 mg of haloperidol, followed by continuous infusion at a rate of 0.1 mg/h for 12 h (total dose per day: 1.7 mg) reported by Wang et al. High-dose haloperidol may be necessary for the primary prevention of delirium, but such treatment may increase the frequency and severity of adverse effects, particularly in vulnerable patients. It has been suggested that haloperidol should ideally be administered to elderly patients at a low dose for a short time [29].

It is also possible that the initiation of preventive administration in the present study may have been too late. In the time-course of the NEECHAM scores after surgery (Fig. 3), the lowest score was noted on postoperative day 1 in both the intervention and control groups, and it slowly improved thereafter. The NEECHAM score may have rapidly decreased immediately after surgery within one postoperative day. The drug was administered early after surgery in the studies by Kaneko et al. and Wang et al., in which an effect was observed. We performed our study on the assumption that the intervention would be administered in general wards of general hospitals, and decided to start haloperidol administration on postoperative day 1 in consideration of the patient safety; however, this timing was not before the decrease in the NEECHAM score was observed. This may have been the reason for the absence of an effect due to the intervention. Although it may have been better to initiate the intervention on the night after the surgery, an intervention immediately after surgery to patient in unstable general conditions is difficult unless it is strictly monitored in the ICU, as reported by Wang et al. When preventive administration is initiated immediately after surgery, the influence of anesthesia remains, and the respiratory and circulatory dynamics are unstable, which may have a negative influence on the postoperative course and increase the possibility of severe adverse events.

It should be noted that we were unable to exclude the possibility of a psychological effect arising from the patients' awareness of having received a haloperidol drip infusion, because this trial was not performed in a double-blinded manner. However, a lack of blinding would introduce potential bias in favor of the haloperidol prophylactic group, which was not observed in the present study. Therefore, the non-double-blinded nature of the study does not appear to have affected the conclusions regarding the treatment.

No severe adverse event corresponding to grade 3 or more that was assumed to be associated with the preventive haloperidol administration was noted. Although high-dose haloperidol administration (from 5 mg to more than 20 mg) was necessary when postoperative delirium developed, no adverse event assumed to be induced by the haloperidol was noted even in these cases, which suggested that haloperidol can be relatively safely administered even during the unstable postoperative period. The safety of haloperidol for postoperative elderly patients has also been confirmed in other studies [13–15]. Combining its high tolerability and low cost, expanding the clinical experience using this antipsychotic drug will broaden the range of applications for other conditions.

Nevertheless, this study suggests that intravenous treatment with low-dose haloperidol is unlikely to be used widely in hospital wards as prophylactic intervention to prevent postoperative delirium in elderly surgical patients who are not orally treatable. High-dose haloperidol may be

effective, but it has been advised that haloperidol should be administered to elderly patients at a low dose to prevent adverse effects [25, 26, 29].

Several studies have shown various non-pharmacological measures that can contribute to reducing the incidence of delirium [30–33]. Inouye et al. [30] proposed a multifactorial intervention that included specific protocols for cognitive impairment, sleep deprivation, immobility, visual impairment, hearing impairment and dehydration by showing significant reductions in the number and duration of episodes of delirium in hospitalized general medicine patients aged 70 or older. Marcantonio et al. [32] reported that geriatric consultation decreased delirium in elderly surgical patients after hip fracture by over one-third. More recently, Colombo et al. [33] showed that the timely use of a re-orientation strategy was correlated with a significantly lower occurrence of delirium in patients admitted to the ICU. In fact, the proportion of patients who developed postoperative delirium in the present study (38 % of all patients) was lower than that reported in our previous study (55 %) [7]. This may have been due to the protocol used in this study, in which patients who developed delirium on the day of surgery or on the following morning were not included. However, alternatively, it could have been because the preceding studies on delirium motivated the participating medical staff and nurses to be more aware of postoperative delirium, thus promoting the care of patients at risk and decreasing the incidence of delirium.

The NEECHAM score was employed to diagnose delirium and to evaluate its severity. A pattern similar to that in our previous report [7] was noted in the present study, which confirmed its reproducibility and usefulness as a score to evaluate delirium. The incidence of postoperative delirium has been shown to vary markedly (10 % to higher than 50 %) among reports [29, 34–38], and this may have been due to the fact that many studies were retrospective, and the definition of the development of delirium has been ambiguous or different among these studies. The DMS-IV [39] and Delirium Rating Scale (DRS) [40] are known as diagnostic criteria that can also be used to evaluate delirium; however, it is difficult for nurses who directly take care of patients with postoperative delirium to judge delirium accurately using the DMS-IV and DRS. On the other hand, the severity and condition of delirium can be simply and objectively determined by general physicians and nurses using the NEECHAM score. In this study, postoperative delirium was continuously and prospectively evaluated in consecutive cases using the NEECHAM score by nurses in direct contact with patients all day at clinical sites. The efficacy of the intervention was judged based on this evaluation, and its reliability should have been high.

Acknowledgments This work was supported by a Research Grant for Longevity Sciences (17C-3, 21-13) from the Ministry of Health, Labour and Welfare and The Research Funding for Longevity Sciences (23-28) from National Center for Geriatrics and Gerontology (NCGG), Japan.

Conflict of interest The authors of this report have no conflicts of interest.

References

1. Committee for Osteoporosis Treatment of The Japanese Orthopaedic. Association. Nationwide survey of hip fractures in Japan. J Orthop Sci. 2004;9(1):1–5.
2. Isobe H, Takasu N, Mizutani M, Kimura W. Management of colorectal cancer in elderly patients over 80 years old (in Japanese). Nippon Ronen Igakkai Zasshi. 2007;44(5):599–605.
3. Blansfield JA, Clark SC, Hofmann MT, Morris JB. Alimentary tract surgery in the nonagenarian: elective vs. emergent operations. J Gastrointest Surg. 2004;8(5):539–42.
4. Amemiya T, Oda K, Ando M, Kawamura T, Kitagawa Y, Fukata S, et al. Activities of daily living and quality of life of elderly patients after elective surgery for gastric and colorectal cancers. Ann Surg. 2007;246(2):222–8.
5. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Fourth Edition. Text Revision (DSM-IV-TR). Washington, DC, American Psychiatric Association, 2000.
6. McCusker J, Cole M, Abrahamowicz M, Primeau F, Belzile E. Delirium predicts 12-month mortality. Arch Intern Med. 2002; 162(4):457–63.
7. Hattori H, Kamiya J, Shimada H, Akiyama H, Yasui A, Fukata S, et al. Assessment of the risk of postoperative delirium in elderly patients using E-PASS and the NEECHAM Confusion Scale. Int J Geriatr Psychiatry. 2009;24:1304–10.
8. Teslyar P, Stock VM, Wilk CM, Camsari U, Ehrenreich MJ, Himelhoch S. Prophylaxis with antipsychotic medication reduces the risk of post-operative delirium in elderly patients: a meta-analysis. Psychosomatics. 2013;54(2):124–31.
9. Hakim SM, Othman AI, Naoum DO. Early treatment with risperidone for subsyndromal delirium after on-pump cardiac surgery in the elderly: a randomized trial. Anesthesiology. 2012; 116(5):987–97.
10. Larsen KA, Kelly SE, Stern TA, Bode RH Jr, Price LL, Hunter DJ, et al. Administration of olanzapine to prevent postoperative delirium in elderly joint-replacement patients: a randomized, controlled trial. Psychosomatics. 2010;51:409–18.
11. Prakanrattana U, Prapajtrakool S. Efficacy of risperidon for prevention of postoperative delirium in cardiac surgery. Anaesth Intensive Care. 2007;35:714–9.
12. Tagarakis GI, Voucharas C, Tsolaki F, Daskalopoulos ME, Papaliagkas V, Parisis C, et al. Ondasetron versus haloperidol for the treatment of postcardiotomy delirium: a prospective, randomized, double-blinded study. J Cardiothorac Surg. 2012;7:25.
13. Kaneko T, Cai J, Ishikura T, Kobayashi M, Naka T, Kaibara N. Prophylactic consecutive administration of haloperidol can reduce the occurrence of postoperative delirium in gastrointestinal surgery. Yonago Ada Med. 1999;42:179–84.
14. Wang W, Li HL, Wang DX, Zhu X, Li SL, Yao GQ, et al. Haloperidol prophylaxis decreases delirium incidence in elderly patients after noncardiac surgery: a randomized controlled trial. Crit Care Med. 2012;40:731–9.
15. Kalisvaart KJ, de Jonghe JF, Bogaards MJ, Vreeswijk R, Egberts TC, Burger BJ, et al. Haloperidol prophylaxis for elderly hip surgery patients at risk fur delirium: A randomized placebo-controlled study. J Am Geriatr Soc. 2005;53:1658–66.
16. Schrader SL, Wellik KE, Demaerschalk BM, Caselli RJ, Woodruff BK, Wingerchuk DM. Adjunctive haloperidol prophylaxis

reduces postoperative delirium severity and duration in at-risk elderly patients. Neurologist. 2008;14(2):134–7.
17. Seitz D, Gill SS. Perioperative haloperidol to prevent postoperative delirium. J Am Geriatr Soc. 2006;54(5):861; author reply 861–3.
18. Wang MD. Perioperative haloperidol usage for delirium management. J Am Geriatr Soc. 2006;54(5):860–1; author reply 861–863.
19. Cockrell JR, Folstein MF. Mini-mental state examination (MMSE). Psychopharmacol Bull. 1988;24:689–92.
20. Anthony JC, LeResche L, Niaz U, von Korff MR, Folstein MF. Limits of the 'Mini-Mental State' as a screening test for dementia and delirium among hospital patients. Psychol Med. 1982;12:397–408.
21. Mahoney FL, Barthel DW. The Barthel. Index. Maryland. State. Mad J. 1965;14:61–5.
22. Neelon VJ, Champagne MT, Carlson JR, Funk SG. The NEECHAM confusion scale: construction, validation, and clinical testing. Nurs Res. 1996;45(6):324–30.
23. Matsushita T, Matsushima E, Maruyama M. Early detection of postoperative delirium and confusion in a surgical ward using the NEECHAM Confusion Scale. Gen Hosp Psychiatry. 2004;26(2):158–63.
24. Schuurmans MJ, Deschamps PI, Markham SW, Shortridge-Baggett LM, Duursma SA. The measurement of delirium: review of scales. Res Theory Nurs Pract. 2003;17(3):207–24.
25. American Psychiatric Association. Practice guideline for the treatment of patients with delirium. Am J Psychiatry. 1999;156(suppl):1–20.
26. Hatta K. Clinical Guideline for the Treatment of Delirium Japanese Society of General Hospital Psychiatry Practice Guidelines 1. Tokyo: Seiwa Shoten; 2005. p. 19–47.
27. Immers HE, Schuurmans MJ, van de Bijl JJ. Recognition of delirium in ICU patients: a diagnostic study of the NEECHAM Confusion Scale in ICU patients. BMC Nurs. 2005;4:7.
28. Forsman A, Ohman R. Pharmacokinetic studied on haloperidol in man. Curr Ther Res Exp. 1976;20:319–36.
29. Cole MG. Delirium in elderly patients. Am J Geriatr Psychiatry. 2004;12:2–7.
30. Inouye SK, Bogardus ST, Carpenter PA, Leo-Summers L, Acampora D, Holford TR, et al. A multicomponent intervention to prevent delirium in hospitalized older patients. N Engl J Med. 1999;340:669–76.
31. Cole MG, Primeau F, McCusker J. Effectiveness of intervention s to prevent delirium in hospitalized patients: a systematic review. CMAJ. 1996;155:1263–8.
32. Marcantonio ER, Flacker JM, Wright RJ, Resnick NM. Reducing delirium after hip fracture: a randomized trial. J Am Geriatr Soc. 2001;49:516–22.
33. Colombo R, Corona A, Praga F, Minari C, Giannotti C, Castelli A, et al. A reorientation strategy for reducing delirium in the critically ill. Results of an interventional study. Minerva Anestesiol. 2012;78(9):1026–33.
34. Demeure MJ, Fain ML. The elderly surgical patient and postoperative delirium. J Am Coll Surg. 2006;203:752–7.
35. Robinson TN, Kiseman U. Postoperative delirium in the elderly: diagnosis and management. Clin Interv Aging. 2008;3:351–5.
36. Olin K, Eriksdotter-Jonhagen M, Jansson A, Herrington MK, Kristiasson M. Permert Jet. Postoperative delirium in elderly patients after major abdominal surgery. Br J Surg. 2005;92(12):1559–64.
37. Goldenberg G, Kiselev P, Bharathan T, Baccash E, Gill L, Madhav V, et al. Predicting postoperative delirium in elderly patients undergoing surgery for hip fracture. Psychogeriatrics. 2006;6:43–8.
38. Santana SF, Wahlund LO, Varli F, Tadeu VI, Eriksdotter JH. Incidence, clinical features and subtypes of delirium in elderly patients treated for hip fractures. Dement Geriatr Cogn Disord. 2005;20(4):231–7.
39. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision (DSM IV-TR).
40. Trzepacz PT, Baker RW, Greenhouse J. A symptom rating scale for delirium. Psychiatry Res. 1988;23:89–97.

VOLUME 32 · NUMBER 18 · JUNE 20 2014

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Randomized Phase III Trial of Erlotinib Versus Docetaxel As Second- or Third-Line Therapy in Patients With Advanced Non–Small-Cell Lung Cancer: Docetaxel and Erlotinib Lung Cancer Trial (DELTA)

Tomoya Kawaguchi, Masahiko Ando, Kazuhiro Asami, Yoshio Okano, Masaaki Fukuda, Hideyuki Nakagawa, Hidenori Ibata, Toshiyuki Kozuki, Takeo Endo, Atsuhisa Tamura, Mitsuhiro Kamimura, Kazuhiro Sakamoto, Michihiro Yoshimi, Yoshifumi Soejima, Yoshio Tomizawa, Shun-ichi Isa, Minoru Takada, Hideo Saka, and Akihito Kubo

See accompanying article on page 1874

A B S T R A C T

Purpose
To investigate the efficacy of erlotinib versus docetaxel in previously treated patients with advanced non–small-cell lung cancer (NSCLC) in an epidermal growth factor receptor (EGFR) –unselected patient population.

Patients and Methods
The primary end point was progression-free survival (PFS). Secondary end points included overall survival (OS), response rate, safety, and analyses on *EGFR* wild-type tumors. Patients with stage IIIB or IV NSCLC, previous treatment with one or two chemotherapy regimens, evaluable or measurable disease, and performance status of 0 to 2 were eligible.

Results
From August 2009 to July 2012, 150 and 151 patients were randomly assigned to erlotinib (150 mg daily) and docetaxel (60 mg/m^2 every 3 weeks), respectively. *EGFR* wild-type NSCLC was present in 109 and 90 patients in the erlotinib and docetaxel groups, respectively. Median PFS for erlotinib versus docetaxel was 2.0 *v* 3.2 months (hazard ratio [HR], 1.22; 95% CI, 0.97 to 1.55; $P = .09$), and median OS was 14.8 *v* 12.2 months (HR, 0.91; 95% CI, 0.68 to 1.22; $P = .53$), respectively. In a subset analysis of *EGFR* wild-type tumors, PFS for erlotinib versus docetaxel was 1.3 *v* 2.9 months (HR, 1.45; 95% CI, 1.09 to 1.94; $P = .01$), and OS was 9.0 *v* 10.1 months (HR, 0.98; 95% CI, 0.69 to 1.39; $P = .91$), respectively.

Conclusion
Erlotinib failed to show an improvement in PFS or OS compared with docetaxel in an EGFR-unselected patient population.

J Clin Oncol 32:1902-1908. © 2014 by American Society of Clinical Oncology

Tomoya Kawaguchi, Kazuhiro Asami, and Shun-ichi Isa, National Hospital Organization Kinki-Chuo Chest Medical Center; Minoru Takada, Koyo Hospital, Osaka; Masahiko Ando, Center for Advanced Medicine and Clinical Research, Nagoya University Hospital; Akihito Kubo, Aichi Medical University School of Medicine, Aichi; Yoshio Okano, National Hospital Organization Kochi Hospital, Kochi; Masaaki Fukuda, National Hospital Organization Nagasaki Medical Center, Nagasaki; Hideyuki Nakagawa, National Hospital Organization Hirosaki Hospital, Hirosaki; Hidenori Ibata, National Hospital Organization Mie Chuo Medical Center, Tsu; Toshiyuki Kozuki, National Hospital Organization Shikoku Cancer Center, Matsuyama; Takeo Endo, National Hospital Organization Mito Medical Center, Mito; Atsuhisa Tamura, National Hospital Organization Tokyo Hospital; Mitsuhiro Kamimura, National Hospital Organization Disaster Medical Center, Tokyo; Kazuhiro Sakamoto, National Hospital Organization Yokohama Medical Center, Yokohama; Michihiro Yoshimi, National Hospital Organization Fukuoka East Medical Center, Fukuoka; Yoshifumi Soejima, National Hospital Organization Ureshino Medical Center, Ureshino; Yoshio Tomizawa, National Hospital Organization Nishigunma Hospital, Gunma; and Hideo Saka, National Hospital Organization Nagoya Medical Center, Nagoya, Japan.

Published online ahead of print at www.jco.org on May 19, 2014

Terms in blue are defined in the glossary, found at the end of this article and online at www.jco.org.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Tomoya Kawaguchi, MD, Department of Internal Medicine, National Hospital Organization Kinki-Chuo Chest Medical Center, 1180 Nagasone-cho, Kita-ku, Sakai, Osaka 591-8555, Japan; e-mail: t-kawaguchi@kch.hosp.go.jp.

0732-183X/14/3218w-1902w/$20.00

DOI: 10.1200/JCO.2013.52.4694

INTRODUCTION

Lung cancer is the leading cause of cancer-related deaths worldwide. Non–small-cell lung cancer (NSCLC) comprises more than 80% of all lung tumors. Approximately two thirds of NSCLCs are diagnosed at advanced stages. The standard first-line treatmentforNSCLC,platinum-baseddoubletchemotherapy, has a response rate of approximately 30%, and the response usually lasts only 4 to 5 months.[1] Second- and third-line chemotherapy has been used to further improve survival. A standard regimen of docetaxel has been established based on results from randomized phase III studies of patients with previously treated advanced NSCLC,[2,3] in whom the median progression-free survival (PFS) in response to docetaxel was 2.0 to 2.5 months.

Epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (TKIs) are active against previously treated NSCLC. Erlotinib, an EGFR-TKI, showed a significant survival benefit in a placebo-controlled phase III trial (BR21), with a median PFS of 2.2 months and hazard ratio (HR) of 0.61.[4] The noninferiority of gefitinib, another EGFR-TKI, to docetaxel in patients with previously treated NSCLC was shown in terms of survival in a global phase III study (Iressa NSCLC Trial Evaluating Response and Survival Versus Taxotere [INTEREST], n = 1,433)[5]

but not in a smaller phase III study in Japan (V15-32, n = 489).[6] A global phase IV study of erlotinib (Tarceva Lung Cancer Survival Treatment [TRUST], n = 6,580) showed a PFS of 3.3 months[7] and a much longer PFS (5.6 months) in an Asian subset.[8] Although both erlotinib and docetaxel are considered standard therapies for previously treated NSCLC, given the favorable survival in erlotinib-treated Asian patients, erlotinib might produce longer PFS than docetaxel in Asian patients with previously treated NSCLC in an EGFR-unselected population.

The Docetaxel and Erlotinib Lung Cancer Trial (DELTA) is a multicenter, open-label, phase III study from Japan. Because gefitinib failed to show noninferiority to docetaxel in the V15-32 trial, we investigated the efficacy and tolerability of erlotinib versus docetaxel as second- or third-line treatment for *EGFR*-unselected patients with NSCLC.

When this study was initiated, EGFR-TKIs were usually used without testing for EGFR mutational status in clinical practice. Then, the pivotal Iressa Pan-Asia Study (IPASS) study showed that gefitinib was superior to carboplatin and paclitaxel in terms of PFS in patients with *EGFR* mutant tumors (HR, 0.48; 95% CI, 0.36 to 0.64), whereas the opposite results were observed in patients with *EGFR* wild-type tumors (HR, 2.85; 95% CI, 2.05 to 3.98) in the first-line setting.[9] Given the advancement of molecular knowledge, we preplanned an analysis to examine the treatment effect in *EGFR* wild-type and *EGFR* mutant disease.

PATIENTS AND METHODS

Patients

This multicenter, open-label, randomized phase III study was sponsored by the National Hospital Organization, an independent administrative agency in Japan. Patients age 20 years or older were eligible if they met the following criteria: pathologically or histologically proven NSCLC with stage IIIB or IV disease (International Union Against Cancer, version 6); previous treatment with one or two chemotherapy regimens, including at least one platinum agent; evaluable or measurable disease by computed tomography (CT) or magnetic resonance imaging; and Eastern Cooperative Oncology Group performance status (PS) of 0 to 2. The main exclusion criteria were previous exposure to EGFR-TKI or docetaxel, symptomatic brain metastasis, and a second active cancer. Patients were also excluded from the study if they had interstitial pneumonia or pulmonary fibrosis detected by chest CT. All enrolled patients provided written informed consent before entering the study. The protocol was approved by the institutional review boards and ethics committees of the National Hospital Organization.

Treatment

Erlotinib (150 mg per day) was administered orally. Docetaxel was administered every 3 weeks as a 1-hour intravenous infusion of 60 mg/m^2 (ie, the approved dose in Japan). Adverse events were monitored and graded according to the Common Terminology Criteria for Adverse Events (version 3.0). Patients received the study treatment until disease progression or intolerable toxicities. Poststudy treatment was given at the discretion of the physician and patient, and cross-over treatment was allowed in this trial.

Assessments

Tumors assessments were performed via CT, spiral CT, or magnetic resonance imaging, and the same methods of measurement were used throughout the study for each patient. PFS was defined as the time from random assignment to the earliest occurrence of disease progression or death from any cause; patients who had not experienced progression or died at data cutoff were censored at the last tumor assessment. Overall survival (OS) was assessed from the date of random assignment to the date of death as the result of any cause, or data were censored at the last date the patient was confirmed to be alive. Tumor response according to RECIST was assessed at baseline, every month for the first 4 months, and every 2 months thereafter. Investigator assessment of best overall tumor response was used for the analysis. Routine laboratory assessments were performed at baseline, every week for the first month, and every 2 to 4 weeks thereafter. *EGFR* mutations were examined in exons 18 to 21 by a highly sensitive polymerase chain reaction (PCR) –based method (ie, the PCR-invader method, peptide nucleic acid–locked nucleic acid PCR clamp method, or cycleave method). These assays were performed in commercial laboratories to which each institute sent the diagnostic tumor samples.[10]

Statistical Analysis

Eligible patients were randomly assigned 1:1 to erlotinib or docetaxel by the minimization method according to sex, performance status, histology, and institution. Efficacy analyses were completed for the intent-to-treat population. Safety analyses were performed for the population who received at least one dose of the trial medication after random assignment. The primary end point was PFS. Secondary end points were OS, response, safety, and analyses on *EGFR* wild-type and mutant tumors. Median PFS was assumed to be 3.5 months and 2.5 months in patients receiving erlotinib and docetaxel, respectively, based on data from previous clinical trials.[2,7,8] The present study was

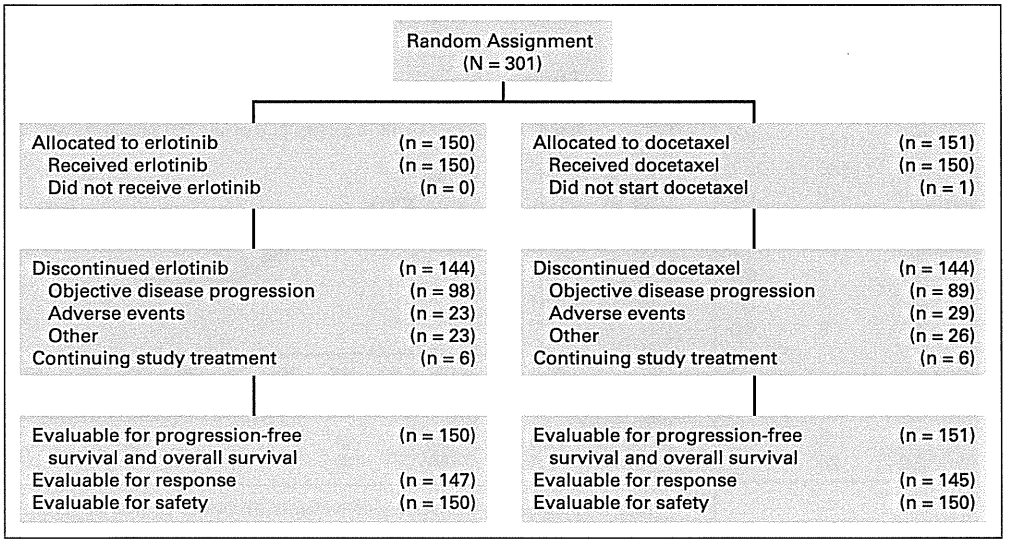

Fig 1. CONSORT diagram.

designed to assess the efficacy of erlotinib versus docetaxel in *EGFR*-unselected patients and to have 80% power to detect a 1-month difference at a two-sided significance level of $P = .05$. A sample size of 300 patients was planned based on these assumptions. Final analysis was planned after 278 events. Survival curves were calculated using the Kaplan-Meier method, and a log-rank test was used to compare treatment groups. The 95% CI of the median survival time was calculated by the method of Brookmeyer and Crowly.[11] Estimates of the treatment effect were expressed as HRs and two-sided 95% CIs from a Cox regression model for erlotinib versus docetaxel.

Subgroup analyses for PFS were performed to explore the potential interaction effect of the treatment groups with sex (male *v* female), PS (0 *v* 1 or 2), stage (IIIB *v* IV), histology (adenocarcinoma *v* other), and smoking status (ever *v* never). Response, toxicity, and patient characteristics were compared between the treatment groups using Fisher's exact test, and age was compared using the Wilcoxon rank sum test. As secondary end points, we performed similar analyses for PFS and OS in patients with *EGFR* wild-type and *EGFR* mutant tumors. To assess the homogeneity of the treatment effect on PFS and OS, an interaction term of treatment and *EGFR* mutation status (wild-type, exon 19 deletion or L858R, or other) was evaluated in the Cox model using the likelihood ratio test. To correct for potential confounding of patient characteristics other than the *EGFR* mutation status in these subgroup analyses, adjusted HRs were also calculated using the Cox regression model, including stratification factors with the exception of institution. Statistical analyses were performed using SAS version 9.3 (SAS Institute, Cary, NC).

Table 1. Patient Demographics and Clinical Characteristics for All Study Patients

Demographic or Clinical Characteristic	Erlotinib (n = 150) No. of Patients	%	Docetaxel (n = 151) No. of Patients	%
Sex				
Female	42	28.0	44	29.1
Male	108	72.0	107	70.9
Age, years				
Median	68		67	
Range	37-82		31-85	
Stage				
IIIB	30	20.0	29	19.2
IV	120	80.0	122	80.8
Performance status				
0	77	51.3	78	51.7
1	67	44.7	67	44.4
2	6	4.0	6	4.0
Smoking status				
Ever-smoker	111	74.0	114	75.8
Never-smoker	39	26.0	37	24.5
Histology				
Adenocarcinoma	104	69.3	103	68.2
Squamous cell carcinoma	29	19.3	32	21.2
Others	17	11.3	16	10.6
First-line treatment	150	100	151	100
Platinum doublet	141	94.0	140	92.7
Platinum doublet + bevacizumab	6	4.0	10	6.6
Other	3	2.0	1	0.7
Second-line treatment	29	19.3	21	13.9
Platinum doublet	19	12.7	9	6.0
Platinum doublet + bevacizumab	3	2.0	3	2.0
Other	7	4.7	9	6.0
EGFR status				
Wild-type	109	72.7	90	59.6
Exon 19 deletion or L858R	21	14.0	30	19.9
Other mutations	2	1.3	3	2.0
Insufficient/not examined	18	12.0	28	18.6

Abbreviation: EGFR, epidermal growth factor receptor.

RESULTS

Patients

From August 2009 to July 2012, 301 patients were enrolled from 41 institutions belonging to the National Hospital Organization. In the intent-to-treat population, 150 and 151 patients were randomly assigned to erlotinib and docetaxel, respectively (Fig 1). The baseline characteristics were well balanced between the treatment groups in terms of age, sex, PS, smoking status, histology, first- and second-line chemotherapy regimens, and *EGFR* status (Table 1).

PFS, OS, and Response Rate in EGFR-Unselected Population

Median PFS time was 2.0 months (95% CI, 1.3 to 2.8 months) for erlotinib and 3.2 months (95% CI, 2.8 to 4.0 months) for docetaxel (Fig 2A), but this difference was not significant (HR, 1.22; 95% CI, 0.97 to 1.55; $P = .09$). At data cutoff (January 17, 2013) with median follow-up of 8.9 months, 141 patients (94.0%) in the erlotinib group and 138 patients (91.4%) in the docetaxel group experienced disease

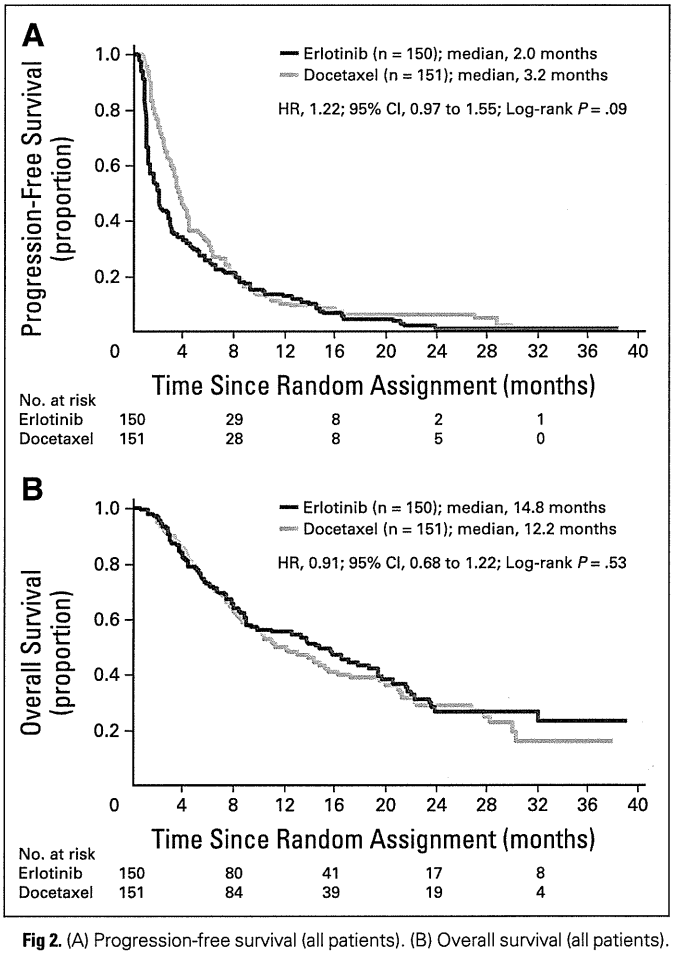

Fig 2. (A) Progression-free survival (all patients). (B) Overall survival (all patients). HR, hazard ratio.

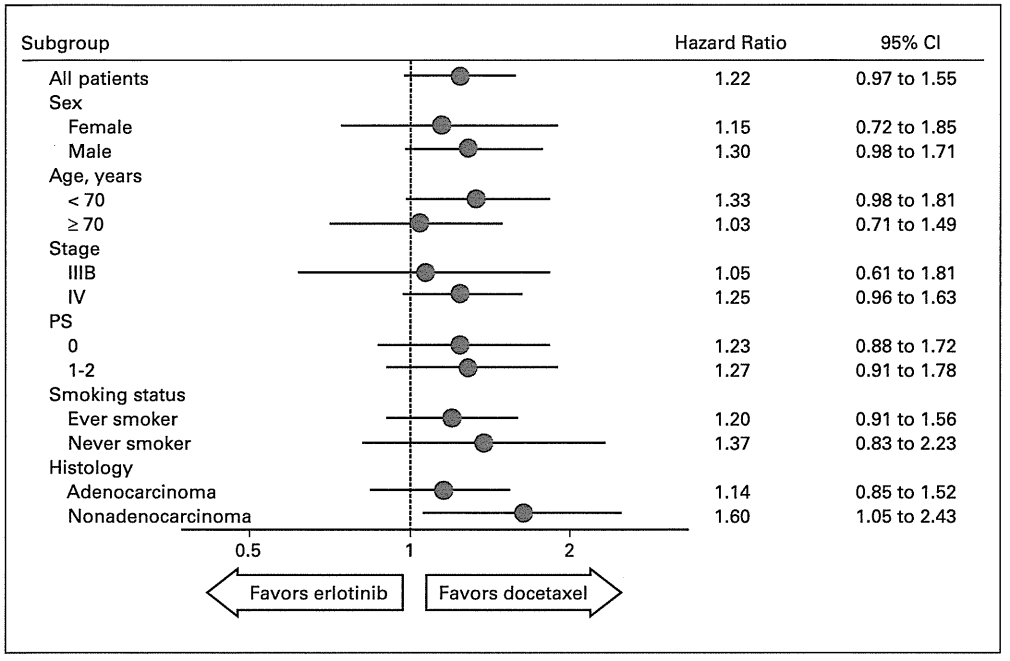

Fig 3. Progression-free survival in clinical subgroups (all patients). PS, performance status.

progression or death. The median OS time was 14.8 months (95% CI, 9.0 to 19.4 months) for erlotinib and 12.2 months (95% CI, 9.0 to 15.5 months) for docetaxel (HR, 0.91; 95% CI, 0.68 to 1.22; $P = .53$; Fig 2B). The number of patients with tumor response was similar in both groups; 25 patients (17.0%; 95% CI, 11.3% to 24.1%) responded in the erlotinib group, and 26 patients (17.9%; 95% CI, 12.1% to 25.2%) responded in the docetaxel group ($P = .88$). A complete response was reported in the erlotinib group in one patient with unknown *EGFR* status. As shown in Figure 3, subgroup analyses for PFS revealed that there was no significant difference between the two drugs, with the exception of nonadenocarcinoma histology (HR, 1.60; 95% CI, 1.05 to 2.43; $P = .03$). All factors numerically favored docetaxel.

PFS, OS, and Response Rate in EGFR *Wild-Type and Mutant Tumors*

EGFR status was determined in 255 (84.7%) of 301 patients, including 199 patients with wild-type *EGFR* NSCLC and 51 patients with active mutant *EGFR* NSCLC. The interaction term between treatment and *EGFR* mutation status was significant for PFS but not for OS ($P = .03$ and $P = .20$, respectively). In patients with *EGFR* wild-type disease, there was no significant difference between the erlotinib and docetaxel groups regarding sex (men and women: 85 and 24 *v* 68 and 22 patients, respectively; $P = .74$), age (median age, 68 *v* 67 years, respectively; $P = .96$), PS (0, 1, and 2: 52, 52, and five *v* 38, 49, and three patients, respectively; $P = .66$), histology (adenocarcinoma and nonadenocarcinoma: 72 and 37 *v* 58 and 32 patients, respectively; $P = .88$), stage (IIIB and IV: 26 and 83 *v* 20 and 70 patients, respectively; $P = .87$), and smoking status (ever-smoker and never-smoker: 87 and 22 *v* 76 and 14 patients, respectively; $P = .46$). In patients with *EGFR* wild-type tumors, the docetaxel group had a significantly longer PFS (2.9 months; 95% CI, 2.1 to 3.3 months) than the erlotinib group (1.3 months; 95% CI, 1.1 to 2.0 months; Fig 4A). A supportive Cox analysis with stratification factors confirmed the significant difference (adjusted HR, 1.57; 95% CI, 1.18 to 2.11; $P < .01$). However, the difference in OS was not statistically significant. The median OS was 9.0 months (95% CI, 7.8 to 14.5 months) in the erlotinib group compared with 10.1 months (95% CI, 7.3 to 12.4 months) in the docetaxel group ($P = .91$; Fig 4B). In terms of tumor response, six patients (5.6%; 95% CI, 2.1% to 11.9%) responded to erlotinib, and 17 patients (20.0%; 95% CI, 12.1% to 30.1%) responded to docetaxel ($P < .01$).

In patients with *EGFR* mutations, median PFS and median OS were longer in the erlotinib group than in the docetaxel group (PFS: 9.3 *v* 7.0 months, respectively; OS: not reached *v* 27.8 months, respectively). However, these differences in PFS (Fig 4C) and OS (Fig 4D) were not statistically significant.

Safety

The safety population included 300 patients: 150 in each group (Table 2). The most common adverse event with erlotinib was rash (92.7%), whereas docetaxel was associated with fatigue (71.3%), nausea (50.0%), and hematologic toxicities. Grade 3 to 4 leukopenia, neutropenia, and febrile neutropenia were significantly more frequent with docetaxel compared with erlotinib (0.7% *v* 64.0%, 0.7% *v* 80.0%, and none *v* 15.3%, respectively; Table 2). Two patients in the erlotinib group died of interstitial lung disease, and one patient in the docetaxel group died as a result of infection.

Poststudy Treatment

The number of patients who received further treatment was similar in the two groups ($P = .22$). Sixty-one patients (42.3%) in the erlotinib group received docetaxel, and 55 patients (37.9%) in the docetaxel group received EGFR-TKIs. Other drugs were administered to 45 patients (31.3%) in the erlotinib group and 41 patients (28.3%) in the docetaxel group. In the unselected population, no difference in OS was observed between the erlotinib and docetaxel arms when comparing patients who went on to receive subsequent chemotherapy (HR, 0.96; 95% CI, 0.62 to 1.49; $P = .84$).

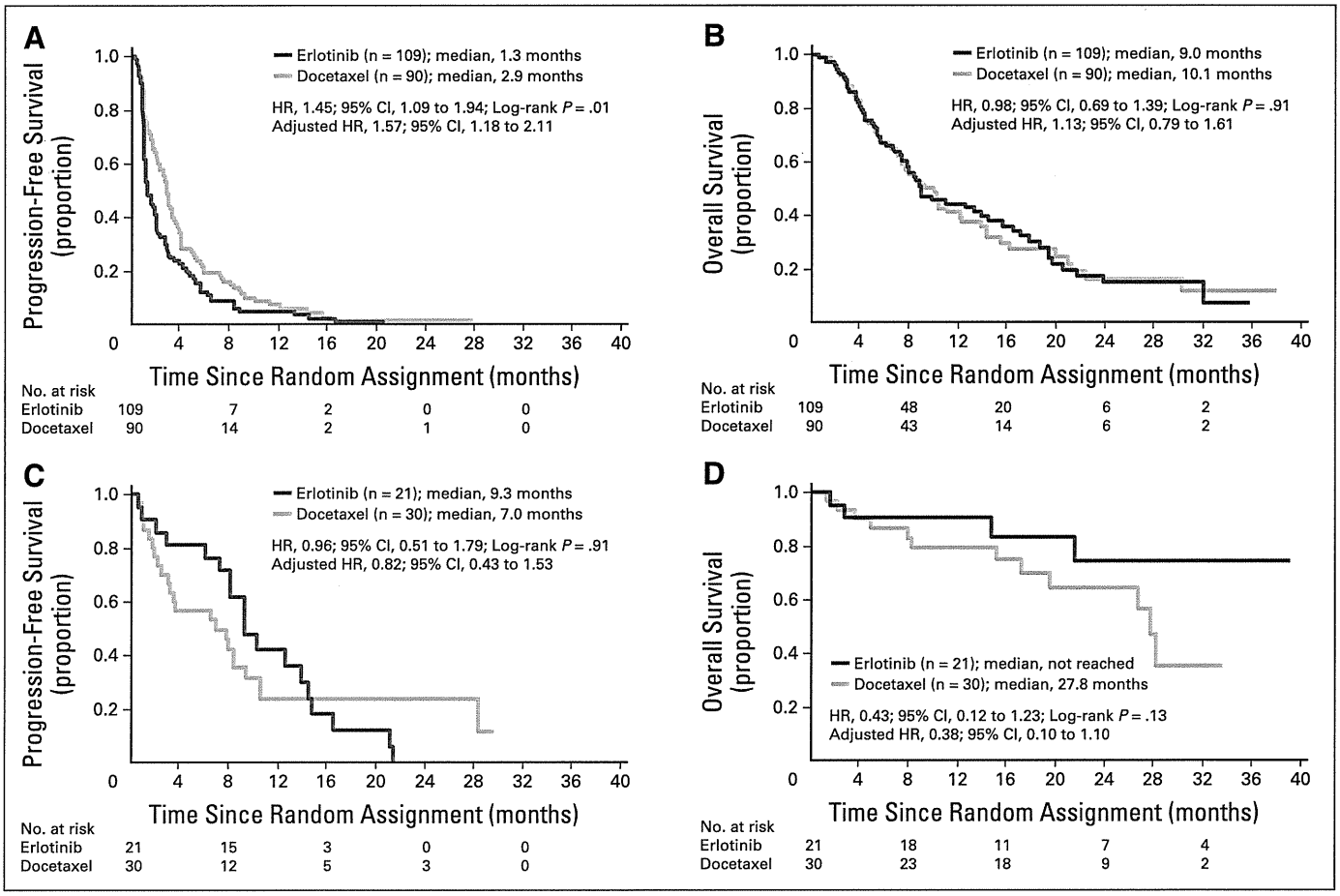

Fig 4. (A) Progression-free survival (PFS) in epidermal growth factor receptor (*EGFR*) wild-type tumors. (B) Overall survival (OS) in *EGFR* wild-type tumors. (C) PFS in *EGFR* mutant tumors (exon 19 deletion or L858R). (D) OS in *EGFR* mutant tumors (exon 19 deletion or L858R). HR, hazard ratio.

Similarly, no difference was observed in the unselected population between the two arms when comparing patients who did not go on to receive subsequent chemotherapy (HR, 1.28; 95% CI, 0.77 to 2.12; $P = .34$). However, patients with *EGFR* wild-type tumors who were treated with docetaxel and did not receive subsequent therapy had a trend toward longer OS when compared with patients treated with erlotinib (HR, 1.79; 95% CI, 0.95 to 3.35; $P = .06$). However, no significant difference in OS was seen between the

Table 2. Common Adverse Events

	All Grades					Grade 3 or 4				
	Erlotinib (n = 150)		Docetaxel (n = 150)			Erlotinib (n = 150)		Docetaxel (n = 150)		
Toxicity	No. of Patients	%	No. of Patients	%	*P*	No. of Patients	%	No. of Patients	%	*P*
Rash	139	92.7	22	14.7	< .01	20	13.3	1	0.7	< .01
Nausea	46	30.7	75	50.0	< .01	3	2.0	5	3.3	.72
Vomiting	13	8.7	25	16.7	.06	1	0.7	0		1.00
Diarrhea	57	38.0	31	20.7	< .01	2	1.3	2	1.3	1.00
Fatigue	80	53.3	107	71.3	< .01	8	5.3	7	4.7	1.00
Anemia	120	80.0	141	94.0	< .01	6	4.0	12	8.0	.22
Thrombocytopenia	31	20.7	48	32.0	.04	0		3	2.0	.245
Leukopenia	19	12.7	140	93.3	< .01	1	0.7	96	64.0	< .01
Neutropenia	15	10.0	136	90.7	< .01	1	0.7	120	80.0	< .01
Neutropenic fever						0		23	15.3	< .01
AST	43	28.7	36	24.0	.43	3	2.0	0		.25
ALT	39	26.0	35	23.3	.69	5	3.3	1	0.7	.21
Pneumonitis	10	6.7	8	5.3	.81	2	1.3	3	2.0	1.00

erlotinib and docetaxel arms in patients who received any subsequent treatment (HR, 0.91; 95% CI, 0.63 to 1.32; $P = .62$).

DISCUSSION

This study showed that there was no significant difference in PFS when comparing erlotinib versus docetaxel as second- or third-line treatment for an *EGFR*-unselected population with NSCLC. In the preplanned subgroup analysis, PFS and response rate were significantly better with docetaxel than erlotinib in *EGFR* wild-type tumors. In contrast, patients with *EGFR* mutant tumors showed longer PFS and OS in the erlotinib group than in the docetaxel group, although these differences did not reach statistical significance, possibly because of the small sample size.

To date, five phase III trials have compared EGFR-TKI and chemotherapy in patients with previously treated and EGFR-unselected NSCLC.[5,6,12-14] INTEREST was the largest study and examined gefitinib versus docetaxel, but there was no significant difference between these two agents in terms of median PFS (2.2 *v* 2.7 months, respectively) and median OS (7.6 *v* 8.0 months, respectively).[5] This trend was also confirmed for Japanese patients in the V15-32 trial.[6] Other drugs examined included erlotinib versus pemetrexed by the Hellenic Oncology Research Group[13] and erlotinib versus docetaxel/pemetrexed in the Tarceva in Treatment of Advanced NSCLC (TITAN) study,[14] and similar results were obtained; there was no difference in PFS and OS between EGFR-TKI and chemotherapy. The findings of DELTA are consistent with the results from these phase III trials in *EGFR*-unselected patients with NSCLC.

Therapy can now be individualized based on the molecular profile of the tumor. Convincing evidence that EGFR-TKIs have marked antitumor activity in patients with activating mutations of exons 19 and 21 of the *EGFR* gene has accumulated.[15,16] This genotyping-guided treatment has been effective in clinical practice. Along with these achievements, the role of EGFR-TKIs in patients with *EGFR* wild-type NSCLC has been discussed.[17] Our prospectively defined analyses included an examination of *EGFR* wild-type NSCLC, revealing 199 patients with wild-type *EGFR* disease (66.1%) among the 255 patients (84.7%) who were assessed for *EGFR* mutations, which is a higher proportion than that assessed in previous studies.[13,14,18] The present analysis showed that docetaxel was superior to erlotinib in terms of PFS in the subset analysis for *EGFR* wild-type NSCLC. To date, three randomized studies have compared EGFR-TKIs and chemotherapy focusing on wild-type *EGFR* tumors.[14,18] However, our data are inconsistent with the subset analyses of the INTEREST[18] and TITAN trials,[14] both of which showed no significant difference in PFS when comparing EGFR-TKIs and chemotherapy. Another recent phase III study, the Tarceva Italian Lung Optimization Trial (TAILOR),[19] in which all the patients had *EGFR* wild-type disease, reported the same results as ours. Because the sample size of the four studies is approximately 200 patients, the discrepancy in PFS among studies might partly be attributable to the methods used for *EGFR* analysis. For example, INTEREST and TITAN used direct sequencing, whereas the TAILOR study used restriction fragment length polymorphism and Sanger sequencing. DELTA adopted highly sensitive PCR-based assays. The TAILOR and DELTA studies used likely more sensitive methods to detect mutations than direct sequencing, particularly for diagnostic tumor samples.[20] The response rates for EGFR-TKI versus docetaxel were 6.6% *v* 9.8%, respectively, in INTEREST; 3.0% *v* 15.5%, respectively, in TAILOR; and 5.6% *v*. 20.0%, respectively, in DELTA (no data available for TITAN). These data support our observations regarding the PFS benefit in the docetaxel group of DELTA.

In contrast to PFS and response rate, there were no differences in OS when comparing EGFR-TKI and chemotherapy in our study as well as in the subset analysis of INTEREST and TITAN. Only the TAILOR study, which did not allow cross-over therapy, showed that docetaxel was better than erlotinib in terms of PFS and OS. In the DELTA study, approximately 40% of patients received cross-over treatments, and other subsequent therapies were similarly delivered in both groups. Therefore, unlike PFS, OS may not be affected by subsequent therapies. In fact, we found a trend toward better OS in the docetaxel group than in the erlotinib group in *EGFR* wild-type patients who received no subsequent chemotherapy in our subset analysis. Given the active drugs available for poststudy chemotherapy that might confer prolonged survival after progression, PFS can be a clinically relevant end point, and further research and discussion are required.[21,22]

The response rate of 20% in the docetaxel arm was higher and hematologic toxicities were more severe compared with the response rate and hematologic toxicities seen in phase III trials in Western countries. There might be some ethnic differences in efficacy and toxicity between white and Asian patients.[23,24] For example, in the Common Arm Trial, which compared clinical outcomes between US and Japanese patients treated with carboplatin and paclitaxel according to identical study design, eligibility criteria, and staging system,[25] the PFS and OS were longer and adverse effects of neutropenia and anemia were more severe in Japanese patients. Although 75 mg/m^2 of docetaxel is more commonly used in Western populations, the absolute response rate and survival in DELTA do not suggest underdosing.

This study has several limitations. First, we failed to detect a significant difference in PFS in the unselected population, which may have been a result of the small sample size. Second, the trial was nonblinded, and the primary end point of PFS was assessed by the individual investigator at each institution. Therefore, caution should be used when comparing our results with those of other studies in which PFS was centrally assessed.

In summary, the present study showed no significant difference in PFS and OS when comparing docetaxel and erlotinib in EGFR-unselected patients with NSCLC. However, docetaxel was superior to erlotinib in terms of PFS and response rate (but not OS) in patients with *EGFR* wild-type disease.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Although all authors completed the disclosure declaration, the following author(s) and/or an author's immediate family member(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a "U" are those for which no compensation was received; those relationships marked with a "C" were compensated. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.
Employment or Leadership Position: None **Consultant or Advisory Role:** None **Stock Ownership:** Masaaki Fukuda, Chugai Pharmaceutical

Honoraria: Tomoya Kawaguchi, Chugai Pharmaceutical, sanofi-aventis; Hideyuki Nakagawa, Chugai Pharmaceutical; Toshiyuki Kozuki, Chugai Pharmaceutical; Yoshio Tomizawa, Chugai Pharmaceutical; Minoru Takada, Chugai Pharmaceutical, sanofi-aventis; Hideo Saka, Chugai Pharmaceutical; Akihito Kubo, Chugai Pharmaceutical, sanofi-aventis
Research Funding: Mitsuhiro Kamimura, Chugai Pharmaceutical; Hideo Saka, Chugai Pharmaceutical **Expert Testimony:** None **Patents, Royalties, and Licenses:** None **Other Remuneration:** None

AUTHOR CONTRIBUTIONS

Conception and design: Tomoya Kawaguchi, Masahiko Ando, Shun-ichi Isa, Minoru Takada, Hideo Saka, Akihito Kubo
Administrative support: Masahiko Ando, Shun-ichi Isa
Provision of study materials or patients: Tomoya Kawaguchi, Kazuhiro Asami, Yoshio Okano, Masaaki Fukuda, Hideyuki Nakagawa, Hidenori Ibata, Toshiyuki Kozuki, Takeo Endo, Atsuhisa Tamura, Mitsuhiro Kamimura, Kazuhiro Sakamoto, Michihiro Yoshimi, Yoshifumi Soejima, Yoshio Tomizawa, Hideo Saka
Collection and assembly of data: Tomoya Kawaguchi, Kazuhiro Asami, Masaaki Fukuda, Hideyuki Nakagawa, Hidenori Ibata, Toshiyuki Kozuki, Takeo Endo, Atsuhisa Tamura, Mitsuhiro Kamimura, Kazuhiro Sakamoto, Michihiro Yoshimi, Yoshifumi Soejima, Yoshio Tomizawa, Hideo Saka, Akihito Kubo
Data analysis and interpretation: Tomoya Kawaguchi, Masahiko Ando, Yoshio Okano, Shun-ichi Isa, Minoru Takada, Hideo Saka, Akihito Kubo
Manuscript writing: All authors
Final approval of manuscript: All authors

REFERENCES

1. Favaretto AG, Pasello G, Magro C: Second and third line treatment in advanced non-small cell lung cancer. Discov Med 8:204-209, 2009
2. Shepherd FA, Dancey J, Ramlau R, et al: Prospective randomized trial of docetaxel versus best supportive care in patients with non-small-cell lung cancer previously treated with platinum-based chemotherapy. J Clin Oncol 18:2095-2103, 2000
3. Fossella FV, DeVore R, Kerr RN, et al: Randomized phase III trial of docetaxel versus vinorelbine or ifosfamide in patients with advanced non-small-cell lung cancer previously treated with platinum-containing chemotherapy regimens: The TAX 320 Non-Small Cell Lung Cancer Study Group. J Clin Oncol 18:2354-2362, 2000
4. Shepherd FA, Rodrigues Pereira J, Ciuleanu T, et al: Erlotinib in previously treated non-small-cell lung cancer. N Engl J Med 353:123-132, 2005
5. Kim ES, Hirsh V, Mok T, et al: Gefitinib versus docetaxel in previously treated non-small-cell lung cancer (INTEREST): A randomised phase III trial. Lancet 372:1809-1818, 2008
6. Maruyama R, Nishiwaki Y, Tamura T, et al: Phase III study, V-15-32, of gefitinib versus docetaxel in previously treated Japanese patients with non-small-cell lung cancer. J Clin Oncol 26:4244-4252, 2008
7. Reck M, van Zandwijk N, Gridelli C, et al: Erlotinib in advanced non-small cell lung cancer: Efficacy and safety findings of the global phase IV Tarceva Lung Cancer Survival Treatment study. J Thorac Oncol 5:1616-1622, 2010
8. Perng RP, Yang CH, Chen YM, et al: High efficacy of erlotinib in Taiwanese NSCLC patients in an expanded access program study previously treated with chemotherapy. Lung Cancer 62:78-84, 2008
9. Mok TS, Wu YL, Thongprasert S, et al: Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med 361:947-957, 2009
10. Goto K, Satouchi M, Ishii G, et al: An evaluation study of EGFR mutation tests utilized for non-small-cell lung cancer in the diagnostic setting. Ann Oncol 23:2914-2919, 2012
11. Brookmeyer R, Crowly J: A confidence interval for the median survival time. Biometrics 38:29-41, 1982
12. Lee DH, Park K, Kim JH, et al: Randomized phase III trial of gefitinib versus docetaxel in non-small cell lung cancer patients who have previously received platinum-based chemotherapy. Clin Cancer Res 16:1307-1314, 2010
13. Karampeazis A, Voutsina A, Souglakos J, et al: Pemetrexed versus erlotinib in pretreated patients with advanced non-small cell lung cancer: A Hellenic Oncology Research Group (HORG) randomized phase 3 study. Cancer 119:2754-2764, 2013
14. Ciuleanu T, Stelmakh L, Cicenas S, et al: Efficacy and safety of erlotinib versus chemotherapy in second-line treatment of patients with advanced, non-small-cell lung cancer with poor prognosis (TITAN): A randomised multicentre, open-label, phase 3 study. Lancet Oncol 13:300-308, 2012
15. Maemondo M, Inoue A, Kobayashi K, et al: Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. N Engl J Med 362:2380-2388, 2010
16. Mitsudomi T, Morita S, Yatabe Y, et al: Gefitinib versus cisplatin plus docetaxel in patients with non-small-cell lung cancer harbouring mutations of the epidermal growth factor receptor (WJTOG3405): An open label, randomised phase 3 trial. Lancet Oncol 11:121-128, 2010
17. Laurie SA, Goss GD: Role of epidermal growth factor receptor inhibitors in epidermal growth factor receptor wild-type non-small-cell lung cancer. J Clin Oncol 31:1061-1069, 2013
18. Douillard JY, Shepherd FA, Hirsh V, et al: Molecular predictors of outcome with gefitinib and docetaxel in previously treated non-small-cell lung cancer: Data from the randomized phase III INTEREST trial. J Clin Oncol 28:744-752, 2010
19. Garassino MC, Martelli O, Broggini M, et al: Erlotinib versus docetaxel as second-line treatment of patients with advanced non-small-cell lung cancer and wild-type EGFR tumours (TAILOR): A randomised controlled trial. Lancet Oncol 14:981-988, 2013
20. Pao W, Ladanyi M: Epidermal growth factor receptor mutation testing in lung cancer: Searching for the ideal method. Clin Cancer Res 13:4954-4955, 2007
21. Broglio KR, Berry DA: Detecting an overall survival benefit that is derived from progression-free survival. J Natl Cancer Inst 101:1642-1649, 2009
22. Booth CM, Eisenhauer EA: Progression-free survival: Meaningful or simply measurable? J Clin Oncol 30:1030-1033, 2012
23. Soo RA, Loh M, Mok TS, et al: Ethnic differences in survival outcome in patients with advanced stage non-small cell lung cancer: Results of a meta-analysis of randomized controlled trials. J Thorac Oncol 6:1030-1038, 2011
24. Hasegawa Y, Kawaguchi T, Kubo A, et al: Ethnic difference in hematological toxicity in patients with non-small cell lung cancer treated with chemotherapy: A pooled analysis on Asian versus non-Asian in phase II and III clinical trials. J Thorac Oncol 6:1881-1888, 2011
25. Gandara DR, Kawaguchi T, Crowley J, et al: Japanese-US common-arm analysis of paclitaxel plus carboplatin in advanced non-small-cell lung cancer: A model for assessing population-related pharmacogenomics. J Clin Oncol 27:3540-3546, 2009

GLOSSARY TERMS

epidermal growth factor receptor (EGFR): also known as HER1. Belongs to a family of receptors (HER2, HER3, HER4 are other members of the family) and binds to the EGF, TGF-α, and other related proteins, leading to the generation of proliferative and survival signals within the cell. It also belongs to the larger family of tyrosine kinase receptors and is generally overexpressed in several solid tumors of epithelial origin.

erlotinib: also known as Tarceva (Genentech, South San Francisco, CA). Erlotinib is a small molecule that inhibits the tyrosine kinase activity of epidermal growth factor receptor/HER1 and has been evaluated extensively in clinical trials in patients with non–small-cell lung cancer, pancreatic cancer, and glioblastoma multiforme.

Acknowledgment

We thank Naoko Akeda for her secretarial assistance.

Supported by a Grant-in-Aid for the Japanese National Hospital Organization Multi-Center Clinical Research for Evidence-Based Medicine, Japan.

Presented in part at the 49th Annual Meeting of the American Society of Clinical Oncology, May 31-June 4, 2013, Chicago, IL.

Cancer Chemother Pharmacol (2014) 74:721–727
DOI 10.1007/s00280-014-2548-z

ORIGINAL ARTICLE

First-line gefitinib therapy for elderly patients with non-small cell lung cancer harboring EGFR mutation: Central Japan Lung Study Group 0901

Kosuke Takahashi · Hiroshi Saito · Yoshinori Hasegawa · Masahiko Ando · Masashi Yamamoto · Eiji Kojima · Yasuteru Sugino · Tomoki Kimura · Fumio Nomura · Tomohiko Ogasawara · Joe Shindoh · Norio Yoshida · Ryujiro Suzuki

Received: 10 January 2014 / Accepted: 25 July 2014 / Published online: 3 August 2014

Abstract
Background The population of elderly patients with lung cancer is increasing worldwide. Although first-line gefitinib is one of the standard treatments for advanced non-small cell lung cancer (NSCLC) harboring epidermal growth factor receptor (EGFR) mutation, few data have been reported regarding gefitinib and elderly patients.
Patients and methods Chemotherapy-naïve patients aged 70 years or older with stage IIIB or IV NSCLC harboring EGFR-activating mutation were enrolled and treated with 250 mg of gefitinib daily until disease progression. The primary end point was response rate, and secondary end points were survival, safety, and quality of life.
Results Twenty patients were enrolled, and the median age was 79.5 years (range 72–90). Overall response rate was 70 % (95 % CI 45.7–88.1 %), and the disease control rate was 90 % (95 % CI 68.3–98.7 %). The median progression-free survival and overall survival time were 10.0 and 26.4 months, respectively. The Functional Assessment of Cancer Therapy-Lung Cancer Subscale (FACT-LCS) scores improved significantly 4 weeks after the initiation of gefitinib ($P = 0.037$) and maintained favorably over a 12-week assessment period. Among the seven items of FACT-LCS, shortness of breath and cough improved

This trial is registered at UMIN-CTR, Number UMIN000001863.

K. Takahashi (✉) · H. Saito
Department of Respiratory Medicine, Aichi Cancer Center Aichi Hospital, 18 Kuriyado Kake-machi, Okazaki, Aichi 444-0011, Japan
e-mail: ktakahashi@acc-aichi.com

Y. Hasegawa
Department of Respiratory Medicine, Nagoya University Graduate School of Medicine, Nagoya, Japan

M. Ando
Center for Medicine and Clinical Research, Nagoya University Hospital, Nagoya, Japan

M. Yamamoto
Department of Respiratory Medicine, Nagoya Ekisaikai Hospital, Nagoya, Japan

E. Kojima
Department of Respiratory Medicine, Komaki Municipal Hospital, Komaki, Japan

Y. Sugino
Department of Respiratory Medicine, Toyota Memorial Hospital, Toyota, Japan

T. Kimura
Department of Respiratory Medicine and Allergy, Tosei General Hospital, Seto, Japan

F. Nomura
Department of Respiratory Medicine, Japanese Red Cross Nagoya Daiichi Hospital, Nagoya, Japan

T. Ogasawara
Department of Respiratory Medicine, Japanese Red Cross Nagoya Daini Hospital, Nagoya, Japan

J. Shindoh
Department of Respiratory Medicine, Ogaki Municipal Hospital, Ogaki, Japan

N. Yoshida
Department of Respiratory Medicine, Kariya Toyota General Hospital, Kariya, Japan

R. Suzuki
Department of Respiratory Medicine, Toyohashi Municipal Hospital, Toyohashi, Japan

significantly after 4 weeks of treatment ($P = 0.046$ and $P = 0.008$, respectively). The most common adverse events were rash and liver dysfunction. Although Grade 1 pneumonitis developed in one patient, no treatment-related death was observed.

Conclusion First-line gefitinib therapy is effective and feasible for elderly patients harboring EGFR mutation, and improves disease-related symptoms, especially pulmonary symptoms like shortness of breath and cough.

Keywords Non-small cell lung cancer · EGFR mutation · Elderly · Gefitinib · Quality of life · First-line treatment

Introduction

Lung cancer is the leading cause of cancer mortality. Non-small cell lung cancer (NSCLC) accounts for 85 % of lung cancer cases, with at least 40 % of the patients at an advanced stage. The population of elderly patients with lung cancer is increasing worldwide. Two-thirds of the lung cancer cases are diagnosed in patients over the age of 65, and the median age at diagnosis is 70 years [1, 2].

Aging is associated with physiologic changes in organ function and altered drug pharmacokinetics. Furthermore, the presence of comorbidities and polypharmacy is frequent in elderly populations. Elderly patients are more likely to experience severe hematologic and non-hematologic toxicity from conventional chemotherapy than their younger counterparts [3]. Before the discovery of driver mutations including epidermal growth factor receptor (EGFR) mutation, single-agent chemotherapy was considered to be a standard of care for elderly patients with advanced NSCLC [4–6]. Although carboplatin and weekly paclitaxel doublet chemotherapy improved overall survival compared with vinorelbine or gemcitabine monotherapy in the IFCT-0501 trial, accompanying toxicity such as Grade 3 or Grade 4 neutropenia, febrile neutropenia, and asthenia was more frequent in the doublet chemotherapy arm [7]. Therefore, investigations of effective treatments with less toxicity are needed for this population.

Gefitinib is an orally administered EGFR tyrosine kinase inhibitor (TKI) that blocks signal transduction pathways implicated in the proliferation and survival of cancer cells. Since EGFR somatic mutation was reported to be strongly related to the response of EGFR-TKI therapy, several studies have demonstrated the efficacy of gefitinib for NSCLC harboring EGFR-activating mutation [8–11]. Two phase III studies comparing gefitinib with platinum doublet chemotherapy as a first-line treatment for NSCLC patients with EGFR mutation showed that the gefitinib group had a higher response and longer progression-free survival than a standard chemotherapy group [12, 13]. However, these studies targeted patients aged 75 years or younger, and few data were available on the efficacy and feasibility of first-line gefitinib therapy for elderly NSCLC patients with EGFR mutation. Therefore, we started our current study of this population. The present study included the assessment of quality of life (QOL) besides the efficacy and feasibility of treatment.

Patients and methods

Patient eligibility

Patients aged 70 years or older with a histologically or cytologically proven diagnosis of non-small cell lung cancer were eligible for this study. Other eligibility criteria included the following: EGFR-activating mutation (either exon 19 deletion or L858R in exon 21); measurable disease; stage IIIB/IV or postoperative recurrence; no prior therapy including chemotherapy or radiotherapy of the primary tumor; Eastern Cooperative Oncology Group (ECOG) performance status of 0–2; an adequate organ function defined as leukocyte count $\geq$3,000/mm^3, platelet count $\geq$100,000/mm^3, hemoglobin $\geq$9.0 g/dl, aspartate aminotransferase and alanine aminotransferase $\leq$100 IU/l, total bilirubin $\leq$1.5 mg/dl, serum creatinine $\leq$1.5 mg/dl, and PaO_2 at rest $\geq$60 mmHg. Patients with any of the following criteria were ineligible: superior vena caval syndrome; history of serious drug allergy; massive pleural or pericardial effusion or ascites that required drainage; interstitial lung disease or pulmonary fibrosis detected by conventional computed tomography of the chest; symptomatic brain metastasis; other concurrent active malignancy; pregnancy, lactation, or other concomitant serious medical conditions. All patients gave written informed consent before enrollment. The study protocol was approved by each institutional review board and was carried out in accordance with the Declaration of Helsinki 1964 (as revised 2000).

Study design and treatment

This was a single-arm, prospective, multicenter, phase II trial. Patients were treated with 250 mg of oral gefitinib daily. Therapy was continued unless there was evidence of disease progression, unacceptable toxicity, or withdrawal of consent. If Grade 3 toxicity other than pneumonitis was observed, gefitinib was discontinued for a maximum of 4 weeks. After the toxicity recovered to the level of Grade 2, gefitinib was given every other day. If toxicity further improved, gefitinib was given daily. If Grade $\geq$1 pneumonitis or Grade 4 toxicity other than pneumonitis was observed, the patient was removed from the study.

Evaluation of response and toxicity

The pretreatment baseline evaluation included a complete medical history and physical examination, complete blood cell count, blood chemistry studies, computed tomography scan of the chest and abdomen, computed tomography or magnetic resonance imaging of the brain, bone scintigraphy or positron emission tomography, arterial blood gas analysis, pulmonary function tests, and electrocardiography. Tumor response was assessed every 2 months during the first year after enrollment and every 3 months between 12 and 18 months. Thereafter, the interval was at the physician's discretion.

The Response Evaluation Criteria in Solid Tumors (RECIST) were used for response assessment [14]. Disease control rate (DCR) was defined as the rate of complete response (CR) plus partial response (PR) plus stable disease (SD). An extramural review was conducted to validate staging and response. Toxicity was evaluated according to the National Cancer Institute Common Terminology Criteria (version 3.0).

Quality of life (QOL) was assessed with the Functional Assessment of Cancer Therapy-Lung Cancer Subscale (FACT-LCS) questionnaire version 4. The maximum attainable score on the FACT-LCS was 28, with which the patient was considered to be asymptomatic. Patients were asked to complete the FACT-LCS questionnaire at the time of enrollment and at 4, 8, and 12 weeks after the initiation of treatment.

Mutational analysis of EGFR

Epidermal growth factor receptor (EGFR) genetic testing methods included either direct sequencing, PCR invader, peptide nucleic acid-locked nucleic acid PCR clamp, or the combination of fragment analysis and the Cycleave method.

Statistical analyses

The primary end point of this study was the response rate. We calculated the sample size based on Simon's two-stage design of the phase II study [15]. Assuming that a response rate of 60 % from eligible patients would indicate potential usefulness, and that a rate of 30 % would be the lower limit of interest (with a power of 0.8 at a one-sided significance level of 0.05), accrual of 17 eligible patients was required. Therefore, we planned to accrue a total of 19 patients, assuming there would be a 10 % dropout rate. The duration of survival was measured from the day of enrollment, and the overall survival curve and progression-free survival curve were calculated according to the method of Kaplan and Meier [16]. Repeated-measures analysis of variance was used to assess the differences in the FACT-LCS between baseline and each point during the treatment. Comparisons of the FACT-LCS scores with the baseline scores were adjusted for multiple comparisons using the Dunnett–Hsu test. The software SAS/Proc Mixed version 9.2 (SAS Institute Inc., Cary, NC) was used for statistical analysis. All comparisons were two-sided, and the statistical significance level was set at $P < 0.05$.

Results

Patient characteristics

Between April 2009 and March 2011, 20 patients were enrolled in this study. Sixteen patients (80 %) were aged 75 years or older, and the median age was 79.5 years (range 72–90 years old) (Table 1). All of the 20 patients had adenocarcinoma, 13 (65 %) were female, two (10 %) had an ECOG performance status of 2, and 12 (60 %) had exon 19 deletion mutations.

Tumor responses and survival

Overall response rate was 70 % (95 % CI 45.7–88.1 %), and the disease control rate was 90 % (95 % CI 68.3–98.7 %) (Table 2). Although the response of one patient who developed pneumonitis was not evaluable, progressive disease was observed in only one patient. The median progression-free survival and overall survival time were 10.0 and 26.4 months, respectively (Figs. 1, 2).

Table 1 Patient characteristics

Characteristics	$N = 20$	(%)
Age, years		
Median (range)	79.5 (72–90)	
Sex		
Male	7	35
Female	13	65
Smoking status		
Never smoker	14	70
Former/current smoker	6	30
ECOG performance status		
0	13	65
1	5	25
2	2	10
Stage		
IIIB	4	20
IV	15	75
Postoperative recurrence	1	5
Type of EGFR mutation		
Exon 19 deletion	12	60
L858R	8	40

ECOG Eastern Cooperative Oncology Group