Although normally present at low or undetectable levels in vivo (Uchida et al., 2009), GM-CSF is critical to the normal functioning of blood neutrophils (Uchida et al., 2007) and alveolar macrophages (Trapnell and Whitsett, 2002). In humans, disruption of GM-CSF signaling by high levels of neutralizing GM-CSF autoantibodies is associated with the development of autoimmune pulmonary alveolar proteinosis (autoimmune PAP), a rare lung disease characterized by surfactant accumulation, respiratory failure, and increased infections (Kitamura et al., 1999; Trapnell et al., 2003; Uchida et al., 2004, 2007). However, GM-CSF autoantibodies are present at low levels in healthy people (Uchida et al., 2009) and levels do not correlate with disease severity in PAP patients (Seymour et al., 2003). This apparent paradox is rectified by the concept of a critical threshold of GM-CSF autoantibodies above which GM-CSF-dependent functions (like surfactant clearance by alveolar macrophages) are reduced to zero regardless of the magnitude of the increase above the critical threshold (Bendtzen et al., 2007). Disruption of GM-CSF receptor function by recessive *CSF2RA* or *CSF2RB* mutations causes hereditary PAP, which is histologically indistinguishable from autoimmune PAP (Martinez-Moczygemba et al., 2008; Suzuki et al., 2008, 2010). PAP also occurs in a heterogeneous group of diseases either as a consequence of an underlying clinical condition presumably affecting the alveolar macrophage function (secondary PAP) (Ishii et al., 2009) or by mutations in genes involved in surfactant production (e.g., *SFTPB, SFTPC, ABCA3, TTF1*) (congenital PAP, and PAP associated with interstitial lung disease)(Nogee, 2010; Whitsett et al., 2004). In genetically modified mice, GM-CSF deficiency causes PAP (Dranoff et al., 1994) while GM-CSF overexpression causes a syndrome of macrophage accumulation, tissue damage, and death (Lang et al., 1987). Increased GM-CSF bioactivity has been implicated in the pathogenesis of rheumatoid arthritis, multiple sclerosis, and other inflammatory and autoimmune diseases and evaluation of GM-CSF antagonist therapy is underway in human clinical trials for multiple clinical indications (Hamilton, 2008). These observations indicate that GM-CSF bioactivity is tightly controlled in healthy people, that loss of tight control is involved in the pathogenesis of multiple diseases and suggest the utility of a clinical test to measure impaired GM-CSF signaling in humans.

CD11b is a cell-adhesion molecule normally present in resting neutrophils within pre-formed granules that is translocated to the cell surface upon neutrophil activation (Graves et al., 1992) including after exposure to an increased concentration of GM-CSF (Condliffe et al., 1998). Previously, we exploited this endogenous priming mechanism in developing the CD11b stimulation index (CD11b-SI), a test to measure impaired GM-CSF signaling in human blood specimens (Uchida et al., 2007). In this assay, fresh, heparinized whole blood is incubated with and without GM-CSF and the mean fluorescence of CD11b on $CD16^{Hi}$ leukocytes is measured by flow cytometry to determine the level of cell-surface CD11b on neutrophils ($CD11b_{Surface}$). The GM-CSF stimulated increase in $CD11b_{Surface}$ (i.e., the stimulation index or SI), is large and readily detected in blood specimens from healthy individuals and zero or severely reduced in patients with autoimmune PAP or hereditary PAP (Uchida et al., 2007, 2009; Suzuki et al., 2008). In the present study, we evaluated and optimized the experimental conditions of the CD11b-SI assay and then validated the test using clinical specimens from patients previously diagnosed with autoimmune PAP and from healthy people.

2. Methods

2.1. Participants

The institutional review board of the Cincinnati Children's Hospital Medical Center (CCHMC) and the University of Tokyo Graduate School of Medicine approved the study. All participants or their legal guardians gave written informed consent, and minors gave assent. Participates included 10 individuals referred for evaluation or treatment of autoimmune PAP diagnosed based on clinical and radiographic findings; an open lung biopsy, transbronchial lung biopsy, or cytologic analysis of bronchoalveolar lavage cells and fluid; and a positive GMAb test performed as described (Uchida et al., 2014). We also studied 34 healthy control subjects who were nonsmokers with no history of major illness and symptom-free at the time of enrollment in the study. Data for some individuals were previously reported (Uchida et al., 2007, 2009; Han et al., 2009).

2.2. Reagents

Primary antibodies included FITC- or PE-conjugated anti-human CD11b, PE-conjugated anti-human CD16 (Miltenyi Biotec K.K., Tokyo, Japan); FITC-rat IgG (isotype control) (abcam, Cambridge, UK); anti-human CD11b monoclonal (BioLegend, San Diego, CA); anti-human phospho-STAT5 (Millipore, Billerica, MA). Secondary detection antibody (anti-rat IgG-HRP(HAF005)) was from R&D systems (Minneapolis, MN) and anti-mouse IgG-HRP was from GE Healthcare (Little Challfont Buckinghamshire, UK). HRP-conjugated anti-actin antibody was from Santa Cruz Biotechnology (Dallas, TX). *E. coli*-derived recombinant human GM-CSF was from ATGen (Seongnam, South Korea). Recombinant cytokines, and other proteins included G-CSF, C5a, Interleukin (IL)-6, IL-8, IL-10, Interferon (IFN)-β, IFN-γ (all from Wako, Osaka, Japan).

2.3. Flow cytometry

Evaluation of neutrophils by flow cytometry was performed as previously described (Uchida et al., 2007). Briefly, 200 μL of whole blood in a 1.5 ml polypropylene tube was incubated (37 °C, 30 minutes) with recombinant human GM-CSF (10 ng/mL) or other cytokines / chemokines, placed on ice immediately. An aliquot (50 μL) of each was mixed with 50 μL of ice-cold PBS containing FITC-conjugated anti-human CD11b antibody (4 μL, undiluted) and PE-conjugated anti-human CD16 antibody (4 μL, undiluted) in a polystyrene round-bottom tube (BD Bioscience, San Jose, CA) and incubated (4 °C, 30 min, in darkness). After adding 2 mL of FACS lysing solution (BD Biosciences, San Jose, CA), incubation (room temp, 10 min), tubes were centrifuged (300 ×g, room temp, 5 min). The cell pellets were re-suspended in 200 μL of PBS and stored in 4 °C in the dark until evaluation by flow cytometry. The amount of cell-surface CD11b on neutrophils ([$CD11b_{Surface}$]) was measured by flow cytometry as mean

fluorescent intensity of FITC on $CD16^{Hi}$ cells defined by gating. The CD11b stimulation index (CD11b-SI) was calculated using the following equation

$$CD11b\text{-}SI = ([CD11b_{Surface}]_{+GM\text{-}CSF} - [CD11b_{Surface}]_{No\ GM\text{-}CSF}) \times 100/[CD11b_{Surface}]_{No\ GM\text{-}CSF}$$

where $[CD11b_{Surface}]_{+GM\text{-}CSF}$ represents $[CD11b_{Surface}]$ after incubation with GM-CSF and $[CD11b_{Surface}]_{No\ GM\text{-}CSF}$ represents $[CD11b_{Surface}]$ after incubation without GM-CSF. In some experiments, cytoplasmic CD11b was stained with PE-CD11b (4 μl, undiluted) using the Intracellular Cytokine Staining Kit (eBioscience, San Diego, CA) as per manufacturer's instructions.

2.4. Western blotting

Heparinized whole blood was incubated with or without GM-CSF (10 ng/mL, 37 °C, 15 min) and red blood cells were lysed with lysing buffer (BD Pharmlyse™, BD Biosciences, San Jose, CA). After washing with ice-cold PBS, cells were extracted with protein extraction buffer (M-PER® #78501) containing protease inhibitor cocktail (0.5% v/v), phosphatase inhibitor cocktail (1% v/v) (all from Thermo Scientific, Rockford, IL), and EDTA (5 mM). Protein extracts were suspended in sample loading buffer, layered onto AnykD gradient gels (Bio-rad, Hercules, CA), separated by gel electrophoresis (100 V, 150 minutes) and transferred (Transblot™, Bio-rad, Hercules, CA) onto PVDF membranes (Immobilon, Merck Millipore, Billerica, MA), per manufacturer's instructions. Membranes were incubated with either murine anti-CD11b monoclonal antibody (10 μl from the vial diluted in 500 μl of immunoblotting buffer (Uchida 2009) (BioLegend, San Diego, CA) or murine anti-phospho- STAT5 antibody (5 μl from the vial diluted in 1000 μl of immunoblotting buffer (Millipore, Billerica, MA)) and detection was done using ECL plus™ (GE healthcare, Little Chalfont, UK). Band intensity of CD11b was quantified by Image Quant TL Analysis Toolbox Ver 7.0 (GE Healthcare, Pittsburgh, PA).

2.5. Statistical analysis

Numerical data were evaluated for a normal distribution using the Kolmogorov-Smirnov test and for equal variance using the Levene median test; parametric data are presented as mean (±SD) and nonparametric data are presented as median and interquartile range (IQR). Statistical comparisons of parametric data were made with Student's t-test for two-group comparisons and with one-way analysis of variance with post hoc analysis according to the Holm-Sidak method for multiple-group comparisons or Dunnett's method for multiple comparisons versus control. Nonparametric data were compared with the use of the Mann-Whitney rank-sum test. Intra-individual percent variation was calculated as the average, for all measured values from each person, of the value minus the mean of all values for an individual divided by the mean and multiplied by 100. Inter-individual percent variation was calculated as the average, for all individuals, of the final value for each individual minus the mean for all individuals divided by the mean for all individuals and multiplied by 100. Receiver operating characteristic (ROC) curve and all other statistical analyses were done using SigmaPlot software, version 12 (Systat Software, SanJose, CA). P values less than 0.05 were considered to indicate statistical significance and are denoted by asterisks ($^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$). All experiments were repeated at least twice with similar results.

3. Results

3.1. Optimization of CD11b-SI components and procedure

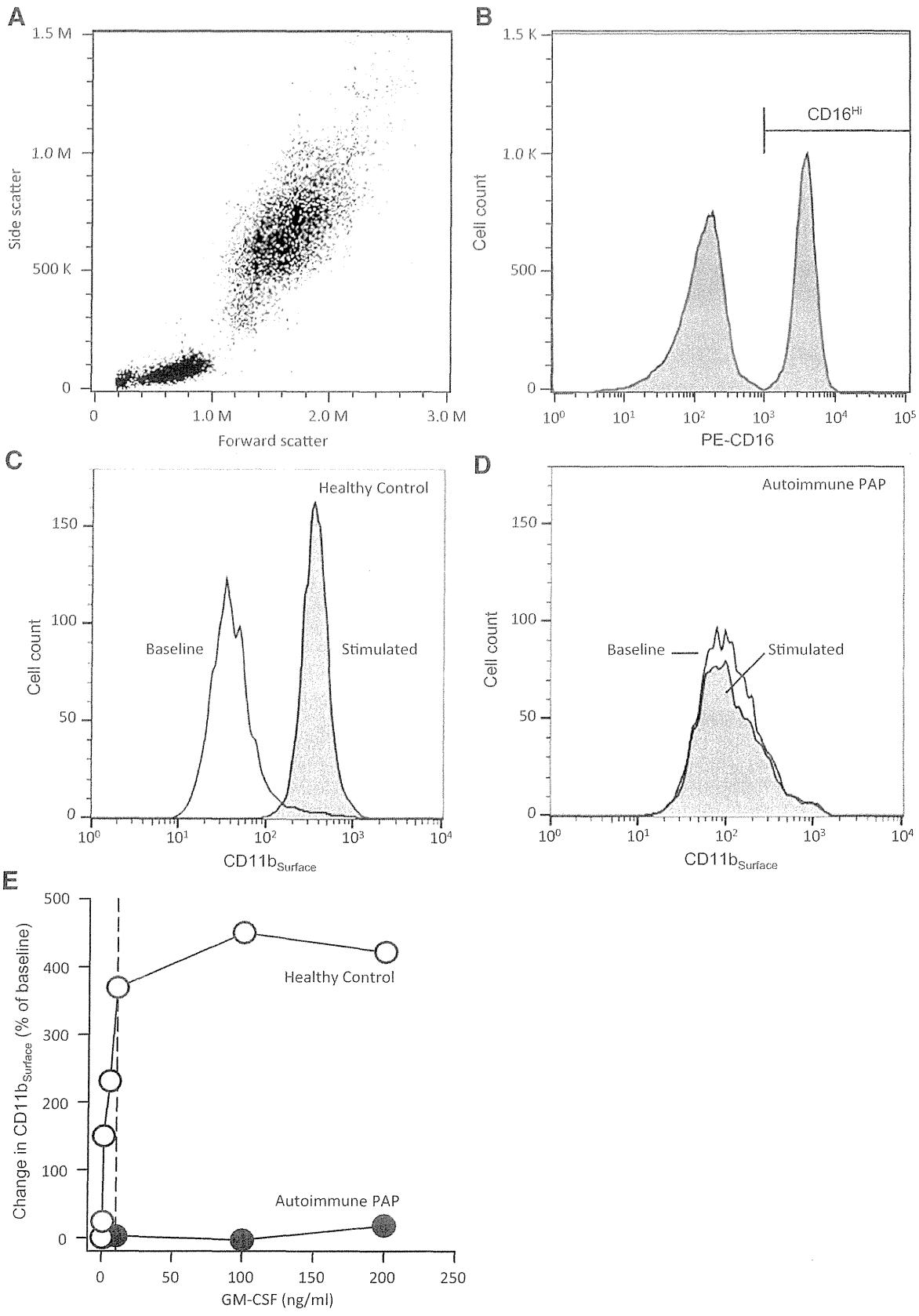
Neutrophils in fresh heparinized whole blood identified by $CD16^{Hi}$ immunostaining (Fig. 1A and B) had detectable levels of cell-surface CD11b ($CD11b_{Surface}$) at baseline (i.e., without stimulation) in both healthy individuals (1C) and autoimmune PAP patients (1D). GM-CSF stimulated a significant increase in $CD11b_{Surface}$ in blood from healthy individuals (Figure C) but not in blood from autoimmune PAP patients (Figure D). The change from baseline $CD1b_{Surface}$ increased with GM-CSF concentration and reached saturation at ~10 ng/mL in healthy people and was severely blunted with little increase in patients with autoimmune PAP (Fig. 1E). To exploit this GM-CSF-stimulated change in $CD11b_{Surface}$ as a measure of impaired GM-CSF signaling in clinical samples, we first optimized experimental conditions to establish a standardized test to measure impaired GM-CSF signaling in whole blood.

The effect of anticoagulant on $CD11b_{Surface}$ was evaluated in fresh whole blood collected into either EDTA- or heparin-containing phlebotomy specimen tubes. Baseline, un-stimulated $CD11b_{Surface}$ was not different among blood specimens collected using EDTA or heparin ($P = 0.174$). In contrast, the GM-CSF-stimulated increase in $CD11b_{Surface}$ was smaller when EDTA was used for phlebotomy compared to samples collected in heparin (Fig. 2).

Blood is frequently collected at room temperature and maintained / evaluated at room temperature, on ice, or at 37 °C, depending on assay requirements. Therefore, we evaluated the effect of temperature on baseline $CD11b_{Surface}$ using fresh whole blood collected at room temperature (~25 °C), and then maintained briefly at different combinations of these three temperatures. When blood was incubated at 25 °C and then at 37 °C (30 min each) or maintained at 37 °C for 60 min, baseline $CD11b_{Surface}$ was unaffected (Fig. 3). However, when blood was incubated at 0 °C and then at 37 °C (30 min each), $CD11b_{Surface}$ increased by 140% (Fig. 3). Consequently, subsequent studies were conducted with blood maintained at room temperature before evaluation.

Next we evaluated potential effects of time before blood specimen evaluation on baseline $CD11b_{Surface}$. The ability of neutrophils to exhibit an increase in $CD11b_{Surface}$ after incubation of heparinized whole blood with 10 ng/mL GM-CSF compared to baseline $CD11b_{Surface}$ (henceforth, change in $CD11b_{Surface}$ (%)) was robust immediately after phlebotomy but markedly diminished after 1 day and declined further thereafter (Fig. 4). In contrast, when blood was stimulated and then stained and fixed immediately after phlebotomy, flow cytometric analysis could be performed for up to seven days with little to no effect on the change in $CD11b_{Surface}$ (Fig. 4).

To further explore the effects of time and temperature on the change in $CD11b_{Surface}$, we evaluated the effects of a short-term delay between phlebotomy and testing for freshly isolated, heparinized whole blood stored at various temperatures. $CD11b_{Surface}$ remained constant for up to 6 hours when


A
1.5 M
1.0 M
500 K
0
Side scatter
0
1.0 M
2.0 M
3.0 M
Forward scatter
B
1.5 K
1.0 K
500
0
Cell count
CD16Hi
10^0
10^1
10^2
10^3
10^4
10^5
PE-CD16
C
150
100
50
0
Cell count
Healthy Control
Baseline
Stimulated
10^0
10^1
10^2
10^3
10^4
CD11bSurface
D
150
100
50
0
Cell count
Autoimmune PAP
Baseline
Stimulated
10^0
10^1
10^2
10^3
10^4
CD11bSurface
E
500
400
300
200
100
0
Change in CD11bSurface (% of baseline)
Healthy Control
Autoimmune PAP
0
50
100
150
200
250
GM-CSF (ng/ml)

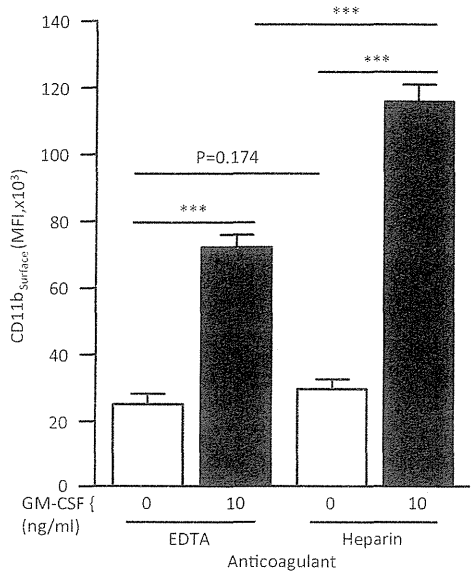


Fig. 2. Effect of anticoagulant on baseline and GM-CSF-stimulated $CD11b_{Surface}$. Fresh blood was collected from healthy people into ethylenediaminetetraacetic acid (EDTA) – or sodium heparin-containing phlebotomy tubes and the mean fluorescence intensity (MFI) of CD11b on $CD16^{Hi}$-gated neutrophils ($CD11b_{Surface}$) was measured after stimulation with 10 ng/ml GM-CSF using the CD11b-SI test as described in the Methods. Bars represent the mean ± SD for n = 3 per condition. Comparisons were made by ANOVA with Dunnett's test; *** = P < 0.001.

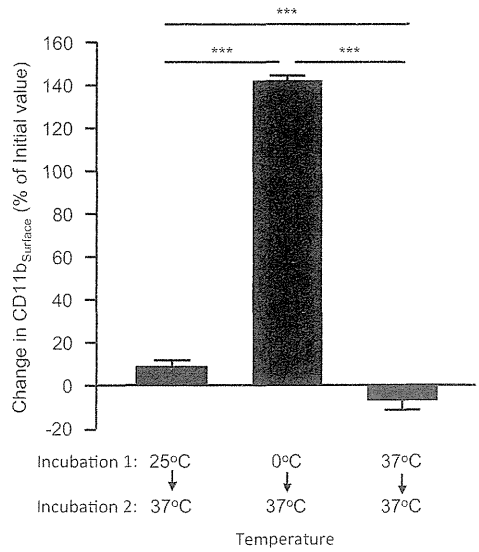


Fig. 3. Effect of change in the temperature at which blood was maintained before testing on the $CD11b_{Surface}$. Fresh, heparinized whole blood was collected from healthy people by phlebotomy and immediately incubated at 0 °C, 25 °C, or 37 °C for 30 min (incubation period 1) followed immediately by incubation at 37 °C for 30 min (incubation period 2) and then $CD11b_{Surface}$ was measured as described in the Methods. For each sample, the $CD11b_{Surface}$ at the end of incubation period 2 was subtracted from the $CD11b_{Surface}$ at the end of incubation period 1, divided by the $CD11b_{Surface}$ at the end of incubation period 1 and multiplied by 100 percent. Positive values indicate an increase in $CD11b_{Surface}$ after incubation period 2 compared to after incubation period 1. Bars represent the mean ± SD for n = 3 per condition. Comparisons were made by ANOVA with Dunnett's test; *** = P < 0.001.

blood was kept on ice but increased progressively when blood was kept at room temperature (and even faster when blood was kept at 37 °C) during this time period (Fig. 5A).

In parallel, the change in $CD11b_{Surface}$ declined over the time period for blood maintained at 25 °C (Fig. 5B).

To better understand the mechanism of the GM-CSF-stimulated increase in $CD11b_{Surface}$, the spatial distribution of CD11b within neutrophils (i.e., in the cytoplasm or on the cell-surface) was evaluated before and after GM-CSF stimulation using several methods. Simultaneous detection of $CD11b_{Surface}$ and cytoplasmic CD11b ($CD11b_{Cytosol}$) in cells double-immunostained with FITC-CD11b or PE-CD11b, respectively (Fig. 6A and B), revealed that $CD11b_{Surface}$ increased by 150 % with GM-CSF stimulation (Fig. 6C). In contrast, $CD11b_{Cytosol}$ decreased by ~11% and $CD11b_{Total}$ was not different (Fig. 6C). CD11bTotal was also unchanged by GM-CSF-stimulation when both $CD11b_{Surface}$ and $CD11b_{Cytosol}$ were measured together using the same detection antibody/chromophore (FITC-anti-human CD11b) (Fig. 6D) or when measured by western blotting (Fig. 6D). Finally, $CD11b_{Surface}$ gradually increased while $CD11b_{Cytosol}$ gradually decreased in fresh whole blood maintained at 25 °C for 6 hours (Fig. 6E). Together, these results indicate that GM-CSF stimulates pre-formed CD11b located within the neutrophil to translocate to the cell surface after stimulation.

The specificity of the change in $CD11b_{Surface}$ was evaluated by comparing the change in $CD11b_{Surface}$ stimulated by GM-CSF to that caused by a wide range of pro- and anti-inflammatory mediators. Compared to baseline, GM-CSF stimulated a greater increase in $CD11b_{Surface}$ than did IL-8, IL-6, G-CSF, interferon (IFN)-β, IL-10, IFN-γ, and complement fragment 5a (Fig. 7).

Taken together, these results indicate that GM-CSF specifically stimulates neutrophils causing a robust and rapid translocation of pre-formed, intracytoplasmic CD11b to the cell surface.

3.2. CD11b-SI test performance

Next, we evaluated the ability of the standardized test to identify impaired GM-CSF signaling in whole blood in patients with autoimmune PAP in whom GM-CSF signaling is known to be impaired by high concentrations of neutralizing GM-CSF autoantibodies and in healthy people in whom GM-CSF signaling in blood leukocytes is present and readily detectable (Uchida et al., 2009). The baseline level of $CD11b_{Surface}$ in autoimmune PAP patients was higher than in healthy individuals (170.2 ± 75.4, 101.4 ± 63.2, n = 10, 25, respectively, p = 0.010; not shown). At the standardized concentration for testing (10 ng/ml), GM-CSF stimulated a mean increase from baseline in $CD11b_{Surface}$ of ~450% in 22 healthy people but no increase in 5 patients with autoimmune PAP (Fig. 8). The difference in GM-CSF-stimulated $CD11b_{Surface}$

Fig. 1. Evaluation of $CD11b_{Surface}$ on the neutrophil in the whole blood: an assay to measure the GM-CSF-stimulated increase in cell-surface CD11b on neutrophils in whole blood. A. Representative leukocyte cytogram. Whole blood was processed and evaluated by flow cytometry as described in the Methods. B. Identification of neutrophils by gating on CD16. Immunostained leukocytes were gated for phycoerythrin (PE)-fluorescence to identify neutrophils as a distinct $CD16^{Hi}$ population. C-D. Quantification of neutrophil cell-surface CD11b ($CD11b_{Surface}$) for a healthy control (C) and a patient with autoimmune PAP (D). Representative histogram of the fluorescence intensity in neutrophils from healthy control. open area – no GM-CSF stimulation; filled area – after GM-CSF stimulation. E. Percent change in $CD11b_{Surface}$ for healthy individual (HC) and patient with autoimmune pulmonary alveolar proteinosis (PAP).

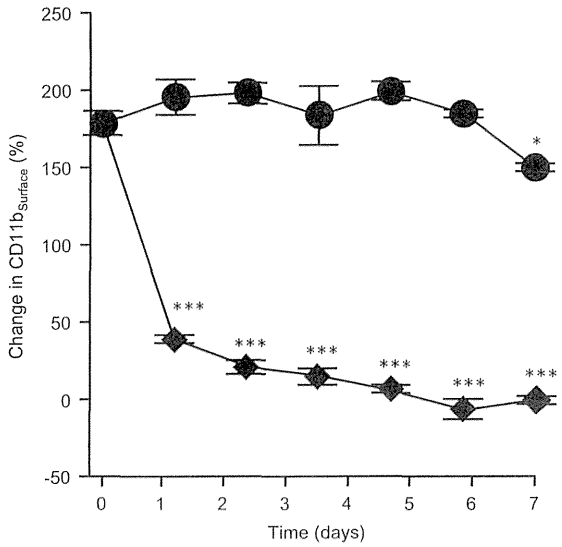


Fig. 4. Effect of time between GM-CSF stimulation/fixation and evaluation by flow cytometry on the change in $CD11b_{Surface}$. Fresh, heparinized blood was collected from healthy people, incubated with or without GM-CSF within 1 hour of phlebotomy, staining with antibodies to $CD11b_{Surface}$, and fixed (closed symbols) or stored at room temperature for various times (open symbols) before stimulation, immunostaining and analysis by flow cytometry as described in the Methods. Symbols represent the mean ± SD for n = 3 per point. Comparisons to the corresponding value on day zero were made by ANOVA with Dunnett's test; * = P < 0.05, *** = P < 0.001.

between healthy people and autoimmune PAP patients was also observed at lower (1 ng/ml) and higher (100 ng/ml) GM-CSF stimulating concentrations, although there was a small increase in the latter (that remained significantly different from healthy controls) (Fig. 8).

We determined CD11b stimulation index or CD11b-SI as change in $CD11b_{Surface}$ at 10 ng/mL of GM-CSF to maximize discrepancies between healthy controls and patients with PAP.

The precision of the CD11b-SI test for measuring impaired GM-CSF signaling in clinical blood specimens was evaluated by measuring the inter-subject and intra-subject variation in fresh heparinized whole blood from nine healthy people (Fig. 9A). The coefficient of variation for repeated measures in the same subject was 5.3 ± 3.5%, which is less than 15% in accordance with FDA guidance criteria for assay precision (Anonymous, 2001) (Fig. 9B). Further, the coefficient of variation for measurements within subject measurements was less than that for measurements between subjects (Fig. 9B). These data indicate the CD11b-SI test is reliable and can be used to measure impaired GM-CSF signaling in clinical blood samples.

3.3. Diagnostic cutoff of the CD11b-SI

To further support the diagnostic use of the CD11b-SI test to diagnose impaired GM-CSF signaling in human clinical specimens, we measured the CD11b-SI in people previously diagnosed with autoimmune PAP and in healthy, asymptomatic people. The CD11b-SI in patients with autoimmune PAP (3.61 [-7.97 - 10.95]; n = 10) was markedly lower than in healthy people (321 [195 - 524]; n = 34) (Fig. 10A). There was a clear separation between the high values in healthy people and the low values in autoimmune PAP patients (Fig. 10B).

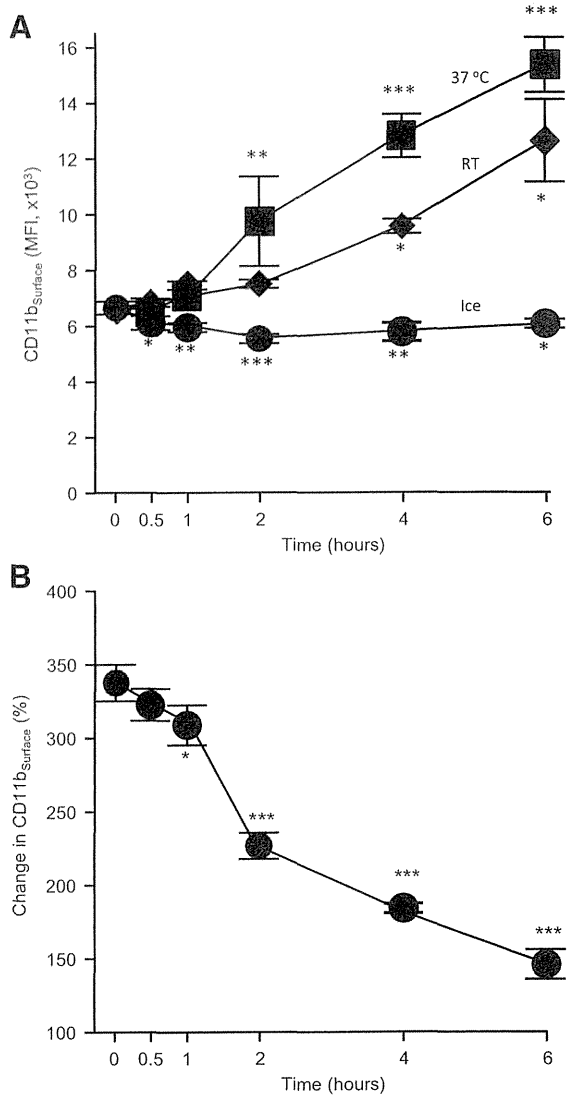


Fig. 5. Effects of temperature of blood after phlebotomy and delay in analysis on change in $CD11b_{Surface}$. A. Effect of temperature on $CD11b_{Surface}$ expression levels. Fresh, heparinized blood was collected and maintained at the indicated temperatures and then CD11b expression levels were evaluated as described in the Methods. B. Effect of a short delay prior to analysis on the change in $CD11b_{Surface}$ stimulated with 10 ng/mL of GM-CSF. Fresh, heparinized blood was kept at 25 °C for various times before evaluation by the change in $CD11b_{Surface}$ as described in the Methods. Symbols represent the mean ± SD for n = 3 per point. Comparisons to the corresponding value at zero hours were made by ANOVA with Dunnett's test; * = P < 0.05, ** = P < 0.01, *** = P < 0.001.

Receiver operating characteristic (ROC) curve analysis for these results demonstrated the sensitivity and specificity to be 100% (Fig. 10C) and identified an optimal cut off threshold value for the difference between normal and abnormal CD11b-SI values (i.e., in healthy people and patients with autoimmune PAP, respectively) to be 112 (Fig. 10D).

4. Discussion

In this study, we optimized the experimental conditions of the CD11b-SI test (Uchida et al., 2007) and then evaluated the

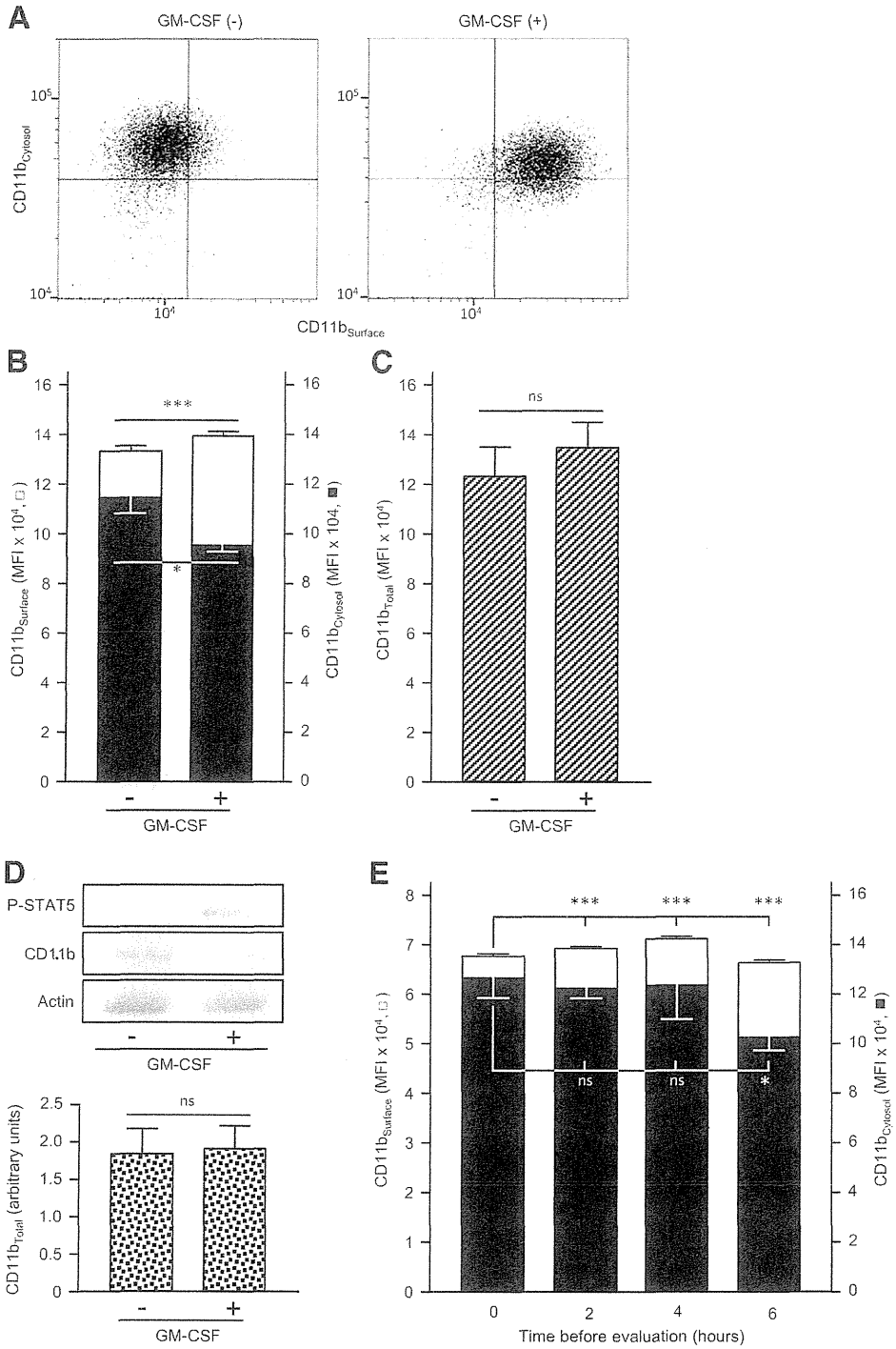


Fig. 6. GM-CSF-stimulated translocation of CD11b to the neutrophil cell-surface. A,B. Fresh, heparinized blood from healthy people was evaluated for the level of CD11b except that cells were first immunostained with FITC-anti-human CD11b to measure $CD11b_{Surface}$ (open bars), washed, permeabilized and stained with PE-anti-human CD11b to measure cytosolic $CD11b_{Cytosol}$ (filled bars) as described in the Methods. A. Representative cytograms of neutrophils stained to show $CD11b_{Surface}$ (horizontal axis) and $CD11b_{Cytosol}$ (vertical axis) with or without GM-CSF stimulation (indicated). B. Quantification of $CD11b_{Surface}$ (open bars) and $CD11b_{Cytosol}$ (closed bars) on neutrophils. Bars represent the mean ± SD for n = 3 per condition. C. Both $CD11b_{Surface}$ and $CD11b_{Cytosol}$ were evaluated with methods described above (legends to Panel A,B) with FITC-anti-human CD11b for both to measure total neutrophil CD11b levels. Bars represent the mean ± SD for n = 3 per condition. D. Quantification of total CD11b by western blotting. Proteins of white blood cell in the whole blood incubated with or without GM-CSF (10 ng/mL) was evaluated by gel electrophoresis and western blotting to quantify CD11b, phosphorylated-STAT5 (as a GM-CSF stimulation control), or actin (as a loading control), and CD11b band intensity quantified by densitometry as described in the Methods. Bars represent mean ± SD for n = 4 determinations per condition. E. Cytosolic CD11b translocation to the neutrophil cell surface with time. Fresh, heparinized blood from healthy people was incubated at room temperature for various times before differential immunostaining to quantify $CD11b_{Surface}$ (open bars) and $CD11b_{Cytosol}$ (filled bars) as described above (legend to Panel A). Bars represent the mean ± SD of n = 3 per condition. Comparisons were made by Student's t test (A-C) or ANOVA (D); * = $P < 0.05$, *** = $P < 0.001$.

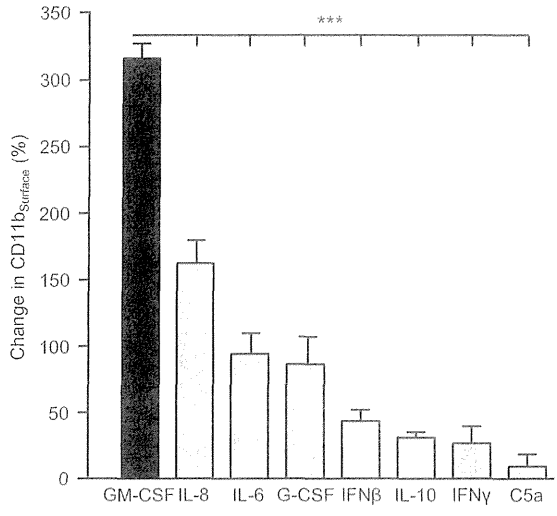


Fig. 7. Capacity of GM-CSF and other neutrophil activators to increase $CD11b_{Surface}$. *Fresh, heparinized blood from healthy people was incubated* without or with GM-CSF or other cytokines, chemokines, and inflammatory molecules including interleukin 8 (IL-8), IL-6, granulocyte-colony stimulating factor (G-CSF), interferon β (IFNβ), IL-10, IFNγ, and complement fragment 5a (C5a), *10 ng/mL each as indicated, and the percent increase in stimulated over* baseline $CD11b_{Surface}$ as described in the Methods. Comparisons were made by ANOVA using the Holm-Sidak method for multiple comparisons; *** = $P < 0.001$.

reliability of the optimized test for detecting impaired GM-CSF signaling in heparinized human blood specimens. In healthy people, GM-CSF rapidly stimulated a robust translocation of pre-formed CD11b to the cell surface of neutrophils while in

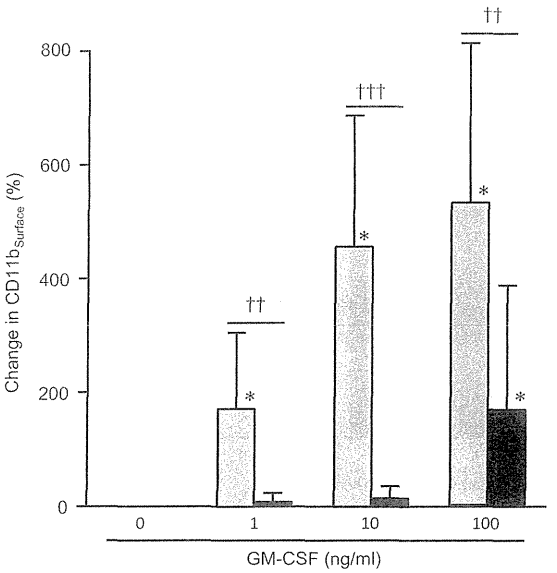


Fig. 8. Detection of impaired GM-CSF signaling in human blood specimens. Fresh, heparinized blood was collected from healthy people (n = 22; gray bars) or patients with autoimmune PAP (n = 5, closed bars) and percent increase of $CD11b_{Surface}$ over baseline as described in the Methods except that different GM-CSF concentrations were used for stimulation (indicated). Comparisons were made by the Mann-Whitney rank sum test with post-hoc analysis using Dunnett's test. comparison of corresponding values for healthy people or PAP patients to the un-stimulated control (no GM-CSF) are indicated; * = $P < 0.05$. Comparison of values for healthy people and PAP patients at each level of GM-CSF stimulation are also indicated; †† = $P < 0.01$, ††† = $P < 0.001$.

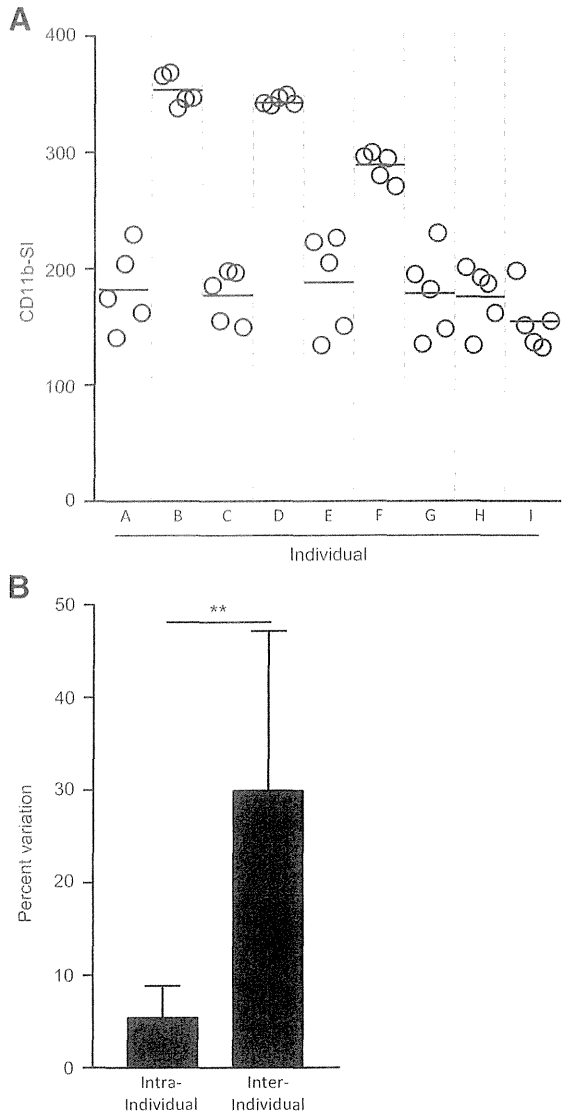


Fig. 9. Precision of the CD11b-SI test measurements in healthy people. A. Fresh, heparinized blood was collected from healthy people (n = 9, A-I) and evaluated repeatedly in independent CD11b-SI test procedures (n = 5 per person). Shown are the individual determinations and mean for each individual (separated by gray dashed lines). Each symbol represents an independent determination. B. Evaluation of the percent variation among repeated determinations of CD11b-SI within the same subject (intra-subject) or between different subjects (inter-subject), calculated as described in the Methods. Comparison was made using Student's t test; ** = $P < 0.01$.

patients with autoimmune PAP, this translocation response was blocked. The assay performed very well in distinguishing impaired GM-CSF signaling in autoimmune PAP patients from normal signaling in healthy people for which the sensitivity and specificity were both 100%.

These results help establish a basis for the routine clinical use of blood testing for the differential diagnosis of PAP. The serum GM-CSF autoantibody ELISA test (GMAb ELISA) used to measure GM-CSF autoantibodies (Schoch et al., 2002) was recently optimized and found to have a sensitivity and specificity of 100% for diagnosis of autoimmune PAP (Uchida et al., 2014). The diagnostic use of the GMAb ELISA was

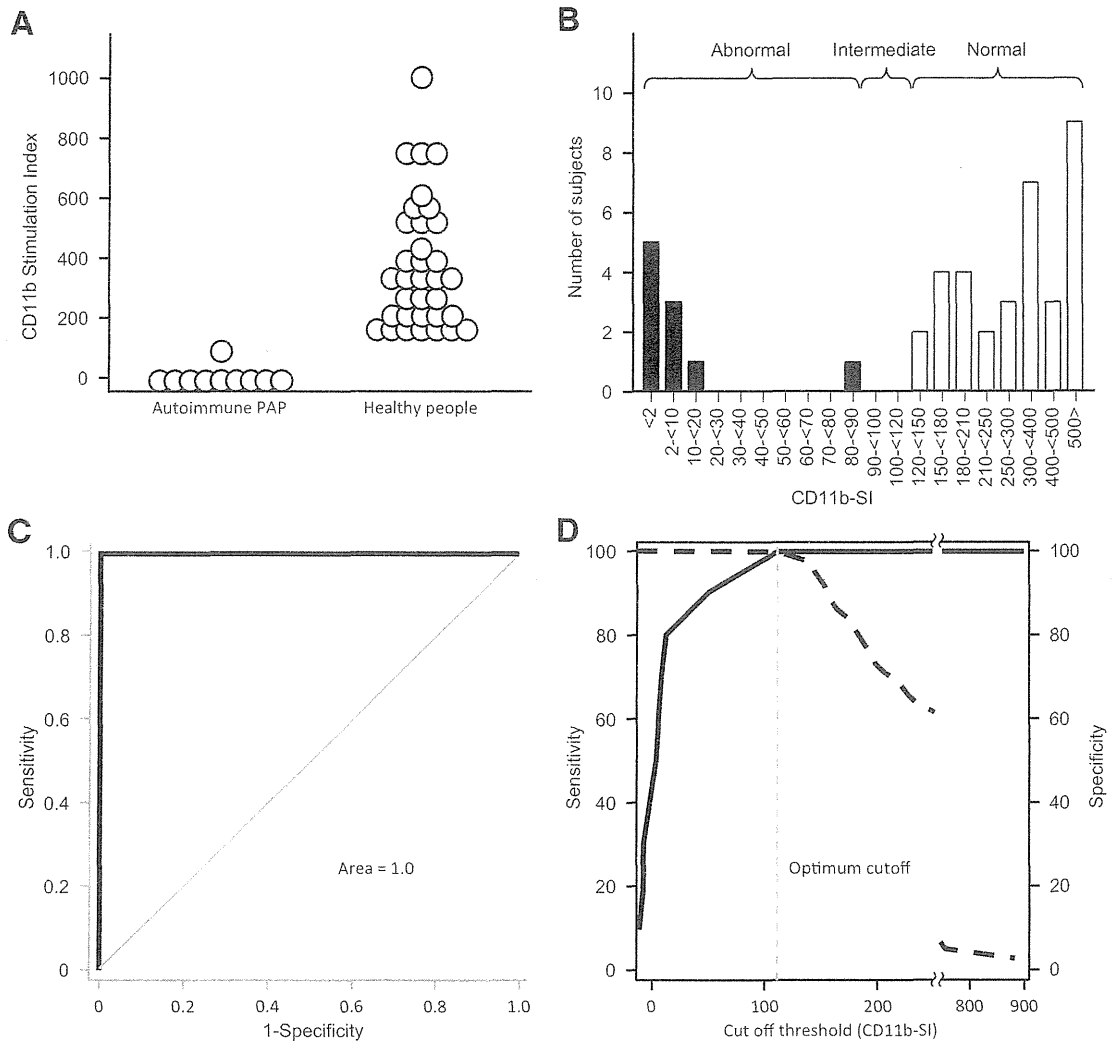


Fig. 10. Measurement and ROC Analysis of CD11b-SI in autoimmune PAP patients and healthy people. A. CD11b-SI test values in autoimmune PAP patients (10) and healthy people (34) measured as described in the Methods. B. Histogram of the frequency distribution of CD11b-SI test values in healthy people (open bars) and autoimmune PAP patients (filled bars). C-D. Receiver operating characteristic (ROC) curve analysis of CD11b-SI results for 10 autoimmune PAP patients and 34 healthy people. Standard ROC analysis was performed to determine the sensitivity and specificity for the data shown in Panel A. The are under the curve was 1.0 (C), and, at a cut off value for CD11b-SI of 112 determined by the software, the sensitivity and specificity of the CD11b-SI were both 100% (D).

supported by the demonstration that GM-CSF autoantibodies are the cause and not an associated epiphenomenon of autoimmune PAP (Sakagami et al., 2009, 2010). Support was also provided by a critical threshold of GM-CSF autoantibodies (Bendtzen et al., 2007) above which the risk of PAP is increased in non-human primates injected with patient-derived GM-CSF autoantibodies (Sakagami et al., 2010) and in humans (Uchida et al., 2009) and by ROC curve analysis confirming the value of the critical threshold (~5 μg/ml) using the optimized GMAb ELISA (Uchida et al., 2014). Notwithstanding, while the latter demonstrated most healthy people and autoimmune PAP patients have GMAb ELISA test results below 3 μg/ml or above 9 μg/ml, respectively, it identified an intermediate concentration range (>3, <9 μg/ml) in which diagnostic utility may be reduced by proximity to the critical threshold. The CD11b-SI test, by measuring impaired GM-CSF signaling in whole blood, can be used to determine the functional significance of an increased GM-CSF autoantibody level when the GMAb ELISA test result is near the critical threshold. Results identified a CD11b-SI of 112 as the optimal cutoff value for distinguishing specimens known to have impaired GM-CSF signaling (e.g., autoimmune PAP) from those of healthy people. However, as a conservative approach to test interpretation, we recommend that test results between 90 and 120 be 'read' as intermediate since data from healthy people or PAP patients to confirm this cutoff were not available in this range. Additional studies may be useful in further evaluating the accuracy, precision, ruggedness of the CD11b-SI for routine clinical use as a diagnostic test. In seven patients with hereditary PAP caused by disruption of GM-CSF receptor function, we found no significant response to GM-CSF stimulation using the CD11b-SI test (Suzuki et al., 2010). Thus, the combination of compatible radiographic findings, a normal GMAb ELISA test result and an abnormal CD11b-SI test result suggest a diagnosis of hereditary PAP. Further, since stimulation at very high concentrations of GM-CSF can stimulate an increase in

$CD11b_{Surface}$ in autoimmune PAP but likely not in hereditary PAP, the use of additional higher concentrations may be useful in the differential diagnosis of hereditary versus autoimmune PAP. However, further studies will be needed to explore and validate such an approach.

The CD11b-SI test may also be useful in other clinical or clinical research settings. For example, since $CD11b_{Surface}$ reflects neutrophil activation, it may be useful in patients with increased GM-CSF bioactivity such as serious infectious or inflammatory diseases (Shakoor and Hamblin, 1992; Palmer and Hamblin, 1993; Cook et al., 2013). Finally, it may be useful as an outcome measure in clinical trials aimed at reducing GM-CSF bioactivity, e.g., in autoimmune or inflammatory diseases (Cook et al., 2013), or the functional evaluation of cross-reacting anti-drug antibodies in individuals treated with recombinant human GM-CSF (Wadhwa et al., 1999). Finally, since this assay measures GM-CSF neutralizing capacity in the serum, it may be useful to monitor autoimmune PAP patients for spontaneous improvement or remission as expected if the GMAb were to fall below critical threshold (Uchida et al., 2009, 2014).

One limitation of the CD11b-SI test is the short time period after phlebotomy during which the test must be conducted in order to obtain useful results. The GM-CSF-stimulated increase in $CD11b_{Surface}$ begins to diminish by 1 hour after phlebotomy, and is markedly blunted by 6 hours, and near zero by 24 hours. This effect is due to an increase in baseline $CD11b_{Surface}$ caused by generation of lysophosphatidylcholines during blood storage, which have potent priming effect on neutrophils (Silliman et al., 1994). Neutrophils undergo apoptosis spontaneously with time after phlebotomy, however, this is not significant at times less than three hours (Homburg et al., 1995; Uchida et al., 2007). These time-dependent effects can be overcome by conducting the GM-CSF stimulation component of the test followed by immunostaining and blood fixation within 1 hour of phlebotomy, after which flow cytometric analysis can be delayed for up to 6 days without compromising the results. Another limitation is the sensitivity of $CD11b_{Surface}$ to changes in the temperature at which blood is kept before testing – i.e., placement on ice followed by warming to either room temp. or 37 °C, which has been reported previously (Shalekoff et al., 1998). In contrast, maintaining blood at 25 or 37 °C gave acceptable results similar to a previous report (Youssef et al., 1995). Use of EDTA as the anticoagulant also reduced the GM-CSF- stimulated increase in $CD11b_{Surface}$, consistent with the contribution of calcium ions to the translocation CD11b to the cell surface (Silliman et al., 2003). Variation in flow cytometer setting-dependent effects (e.g., variable laser output and machine settings) is a limitation (not evaluated here) that could significantly affect test results from day to day, especially if results are reported as a continuous variable corresponding to the numeric value of the CD11b-SI and interpreted based on normal and abnormal ranges of the test result. Finally, the magnitude of the GM-CSF-stimulated $CD11b_{Surface}$ (Fig. 8) and the ability to detect impaired GM-CSF signaling in blood specimens from autoimmune PAP patients (Bendtzen et al., 2007) varied with the concentration of GM-CSF used for stimulation. Thus, $CD11b_{Surface}$ measurement is highly sensitive to pre-analytical conditions of blood preparation and storage as noted in prior efforts to standardize the measurement of neutrophil $CD11b_{Surface}$ (Latger-Cannard et al., 2004).

Our approach to these limitations has been to standardize/control the CD11b-SI assay with respect to each variable. In our standardized CD11b-SI test, blood is collected into sodium heparin- containing phlebotomy tubes, kept at room temperature (~25 °C), stimulated with GM-CSF at a standard concentration (10 ng/ml) within 1 hour after phlebotomy, stained and immediately fixed. Blood from a healthy donor is collected and evaluated in parallel to control for day-to-day, machine, and inter-operator variability. Samples are then subjected to flow cytometric analysis as soon as possible usually in <24 hours, the data are expressed as a percent increase from baseline in $CD11b_{Surface}$ of the GM-CSF-stimulated sample. The test result is reported as a dichotomous variable - either "positive" or "negative" for GM-CSF signaling. These conditions have permitted the CD11b-SI test to be performed rapidly and reliably using whole blood specimens without any purification steps in the context of basic or clinical research. Because neutrophils in autoimmune PAP patients have normal ultrastructure, express expected phenotypic markers, and increase $CD11b_{Surface}$ after removal of GM-CSF autoantibodies, impaired CD11b-SI test results are interpreted as demonstrating the presence of functional GM-CSF autoantibodies (or the neutralizing capacity of whole blood) (Uchida et al., 2007, 2009). Consequently, we have used the CD11b-SI to evaluate impaired GM-CSF signaling in autoimmune PAP patients (Uchida et al., 2007, 2009), animal autoimmune PAP model with non-human primates induced by passive transfer of GM-CSF autoantibodies purified from patients (Sakagami et al., 2009, 2010), and other inflammatory diseases (Han et al., 2009).

Acknowledgements

We thank Carrie Stevens for help with collection of human blood specimens. This work was supported by grants from the National Heart Lung and Blood Institute (R01 HL085453; B.C.T.), National Center for Research Resources (U54 RR0198498, B.C.T. and K.U.), the National Institute for Health and Human Development, Japan Society for the Promotion of Science (A232490720001, B24390364, Y.Y. and K.U.), and Japan Ministry of Health Labor and Welfare (H24-Nanchitou(Nanchi)-Ippan-035, H24-Rinkensui-Ippan-003, K.U.)

References

Anonymous, 2001. Guidance for industry bioanalytical method validation. Guidence Documents, vol. 2012. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Veterinary Medicine (CVM), Bethesda.

Bendtzen, K., Svenson, M., Hansen, M.B., Busch, T., Bercker, S., Kaisers, U., Uchida, K., Beck, D.C., Trapnell, B.C., 2007. GM-CSF autoantibodies in pulmonary alveolar proteinosis. N. Engl. J. Med. 356, 2001.

Condliffe, A.M., Kitchen, E., Chilvers, E.R., 1998. Neutrophil priming: pathophysiological consequences and underlying mechanisms. Clin. Sci. (Lond.) 94, 461.

Cook, A.D., Pobjoy, J., Sarros, S., Steidl, S., Durr, M., Lacey, D.C., Hamilton, J.A., 2013. Granulocyte-macrophage colony-stimulating factor is a key mediator in inflammatory and arthritic pain. Ann. Rheum. Dis. 72, 265.

Dranoff, G., Crawford, A.D., Sadelain, M., Ream, B., Rashid, A., Bronson, R.T., Dickersin, G.R., Bachurski, C.J., Mark, E.L., Whitsett, J.A., et al., 1994. Involvement of granulocyte-macrophage colony-stimulating factor in pulmonary homeostasis. Science 264, 713.

Graves, V., Gabig, T., McCarthy, L., Strour, E.F., Leemhuis, T., English, D., 1992. Simultaneous mobilization of Mac-1 (CD11b/CD18) and formyl peptide chemoattractant receptors in human neutrophils. Blood 80, 776.

Hamilton, J.A., 2008. Colony-stimulating factors in inflammation and autoimmunity. Nat. Rev. Immunol. 8, 533.

Han, X., Uchida, K., Jurickova, I., Koch, D., Willson, T., Samson, C., Bonkowski, E., Trauernicht, A., Kim, M.O., Tomer, G., Dubinsky, M., Plevy, S., Kugathsan, S., Trapnell, B.C., Denson, L.A., 2009. Granulocyte-macrophage colony-stimulating factor autoantibodies in murine ileitis and progressive ileal Crohn's disease. Gastroenterology 136 (1261–71), e1.

Homburg, C.H., de Haas, M., von dem Borne, A.E., Verhoeven, A.J., Reutelingsperger, C.P., Roos, D., 1995. Human neutrophils lose their surface FcγRIII and acquire annexin V binding sites during apoptosis in vitro. Blood 85, 532.

Ishii, H., Trapnell, B.C., Tazawa, R., Inoue, Y., Akira, M., Kogure, Y., Tomii, K., Takada, T., Hojo, M., Ichiwata, T., Goto, H., Nakata, K., 2009. Comparative study of high-resolution CT findings between autoimmune and secondary pulmonary alveolar proteinosis. Chest 136, 1348.

Kitamura, T., Tanaka, N., Watanabe, J., Uchida, Kanegasaki, S., Yamada, Y., Nakata, K., 1999. Idiopathic pulmonary alveolar proteinosis as an autoimmune disease with neutralizing antibody against granulocyte/macrophage colony-stimulating factor. J. Exp. Med. 190, 875.

Lang, R.A., Metcalf, D., Cuthbertson, R.A., Lyons, I., Stanley, E., Kelso, A., Kannourakis, G., Williamson, D.J., Klintworth, G.K., Gonda, T.J., et al., 1987. Transgenic mice expressing a hemopoietic growth factor gene (GM-CSF) develop accumulations of macrophages, blindness, and a fatal syndrome of tissue damage. Cell 51, 675.

Latger-Cannard, V., Besson, I., Doco-Lecompte, T., Lecompte, T., 2004. A standardized procedure for quantitation of CD11b on polymorphonuclear neutrophil by flow cytometry: potential application in infectious diseases. Clin. Lab. Haematol. 26, 177.

Lieschke, G.J., Burgess, A.W., 1992. Granulocyte colony-stimulating factor and granulocyte-macrophage colony-stimulating factor (1). N. Engl. J. Med. 327, 28.

Martinez-Moczygemba, M., Doan, M.L., Elidemir, O., Fan, L.L., Cheung, S.W., Lei, J.T., Moore, J.P., Tavana, G., Lewis, L.R., Zhu, Y., Muzny, D.M., Gibbs, R.A., Huston, D.P., 2008. Pulmonary alveolar proteinosis caused by deletion of the GM-CSFRalpha gene in the X chromosome pseudoautosomal region 1. J. Exp. Med. 205, 2711.

Nogee, L.M., 2010. Genetic basis of children's interstitial lung disease. Pediatr. Allergy Immunol. Pulmonol. 23, 15.

Palmer, S., Hamblin, A.S., 1993. Increased CD11/CD18 expression on the peripheral blood leucocytes of patients with HIV disease: relationship to disease severity. Clin. Exp. Immunol. 93, 344.

Sakagami, T., Uchida, K., Suzuki, T., Carey, B.C., Wood, R.E., Wert, S.E., Whitsett, J.A., Trapnell, B.C., Luisetti, M., 2009. Human GM-CSF autoantibodies and reproduction of pulmonary alveolar proteinosis. N. Engl. J. Med. 361, 2679.

Sakagami, T., Beck, D., Uchida, K., Suzuki, T., Carey, B.C., Nakata, K., Keller, G., Wood, R.E., Wert, S.E., Ikegami, M., Whitsett, J.A., Luisetti, M., Davies, S., Krischer, J.P., Brody, A., Ryckman, F., Trapnell, B.C., 2010. Patient-derived granulocyte/macrophage colony-stimulating factor autoantibodies reproduce pulmonary alveolar proteinosis in nonhuman primates. Am. J. Respir. Crit. Care Med. 182, 49.

Schoch, O.D., Schanz, U., Koller, M., Nakata, K., Seymour, J.F., Russi, E.W., Boehler, A., 2002. BAL findings in a patient with pulmonary alveolar proteinosis successfully treated with GM-CSF. Thorax 57, 277.

Seymour, J.F., Doyle, I.R., Nakata, K., Presneill, J.J., Schoch, O.D., Hamano, E., Uchida, K., Fisher, R., Dunn, A.R., 2003. Relationship of anti-GM-CSF antibody concentration, surfactant protein A and B levels, and serum LDH to pulmonary parameters and response to GM-CSF therapy in patients with idiopathic alveolar proteinosis. Thorax 58, 252.

Shakoor, Z., Hamblin, A.S., 1992. Increased CD11/CD18 expression on peripheral blood leucocytes of patients with sarcoidosis. Clin. Exp. Immunol. 90, 99.

Shalekoff, S., Page-Shipp, L., Tiemessen, C.T., 1998. Effects of anticoagulants and temperature on expression of activation markers CD11b and HLA-DR on human leukocytes. Clin. Diagn. Lab. Immunol. 5, 695.

Silliman, C.C., Clay, K.L., Thurman, G.W., Johnson, C.A., Ambruso, D.R., 1994. Partial characterization of lipids that develop during the routine storage of blood and prime the neutrophil NADPH oxidase. J. Lab. Clin. Med. 124, 684.

Silliman, C.C., Elzi, D.J., Ambruso, D.R., Musters, R.J., Hamiel, C., Harbeck, R.J., Paterson, A.J., Bjornsen, A.J., Wyman, T.H., Kelher, M., England, K.M., McLaughlin-Malaxecheberria, N., Barnett, C.C., Aiboshi, J., Bannerjee, A., 2003. Lysophosphatidylcholines prime the NADPH oxidase and stimulate multiple neutrophil functions through changes in cytosolic calcium. J. Leukoc. Biol. 73, 511.

Suzuki, T., Sakagami, T., Rubin, B.K., Nogee, L.M., Wood, R.E., Zimmerman, S.L., Smolarek, T., Dishop, M.K., Wert, S.E., Whitsett, J.A., Grabowski, G., Carey, B.C., Stevens, C., van der Loo, J.C., Trapnell, B.C., 2008. Familial pulmonary alveolar proteinosis caused by mutations in CSF2RA. J. Exp. Med. 205, 2703.

Suzuki, T., Sakagami, T., Young, L.R., Carey, B.C., Wood, R.E., Luisetti, M., Wert, S.E., Rubin, B.K., Kevill, K., Chalk, C., Whitsett, J.A., Stevens, C., Nogee, L.M., Campo, I., Trapnell, B.C., 2010. Hereditary pulmonary alveolar proteinosis: pathogenesis, presentation, diagnosis, and therapy. Am. J. Respir. Crit. Care Med. 182, 1292.

Trapnell, B.C., Whitsett, J.A., 2002. Gm-CSF regulates pulmonary surfactant homeostasis and alveolar macrophage-mediated innate host defense. Annu. Rev. Physiol. 64, 775.

Trapnell, B.C., Whitsett, J.A., Nakata, K., 2003. Pulmonary alveolar proteinosis. N. Engl. J. Med. 349, 2527.

Uchida, K., Nakata, K., Trapnell, B.C., Terakawa, T., Hamano, E., Mikami, A., Matsushita, I., Seymour, J.F., Oh-Eda, M., Ishige, I., Eishi, Y., Kitamura, T., Yamada, Y., Hanaoka, K., Keicho, N., 2004. High-affinity autoantibodies specifically eliminate granulocyte-macrophage colony-stimulating factor activity in the lungs of patients with idiopathic pulmonary alveolar proteinosis. Blood 103, 1089.

Uchida, K., Beck, D.C., Yamamoto, T., Berclaz, P.Y., Abe, S., Staudt, M.K., Carey, B.C., Filippi, M.D., Wert, S.E., Denson, L.A., Puchalski, J.T., Hauck, D.M., Trapnell, B. C., 2007. GM-CSF autoantibodies and neutrophil dysfunction in pulmonary alveolar proteinosis. N. Engl. J. Med. 356, 567.

Uchida, K., Nakata, K., Suzuki, T., Luisetti, M., Watanabe, M., Koch, D.E., Stevens, C.A., Beck, D.C., Denson, L.A., Carey, B.C., Keicho, N., Krischer, J.P., Yamada, Y., Trapnell, B.C., 2009. Granulocyte/macrophage-colony-stimulating factor autoantibodies and myeloid cell immune functions in healthy subjects. Blood 113, 2547.

Uchida, K., Nakata, K., Carey, B., Chalk, C., Suzuki, T., Sakagami, T., Koch, D.E., Stevens, C., Inoue, Y., Yamada, Y., Trapnell, B.C., 2014. Standardized serum GM-CSF autoantibody testing for the routine clinical diagnosis of autoimmune pulmonary alveolar proteinosis. J. Immunol. Methods 402, 57.

Wadhwa, M., Skog, A.L., Bird, C., Ragnhammar, P., Lilljefors, M., Gaines-Das, R., Mellstedt, H., Thorpe, R., 1999. Immunogenicity of granulocyte-macrophage colony-stimulating factor (GM-CSF) products in patients undergoing combination therapy with GM-CSF. Clin. Cancer Res. 5, 1353.

Whitsett, J.A., Wert, S.E., Trapnell, B.C., 2004. Genetic disorders influencing lung formation and function at birth. Hum. Mol. Genet. 13, R207.

Youssef, P.P., Mantzioris, B.X., Roberts-Thomson, P.J., Ahern, M.J., Smith, M.D., 1995. Effects of ex vivo manipulation on the expression of cell adhesion molecules on neutrophils. J. Immunol. Methods 186, 217.

Published Ahead of Print on April 20, 2015 as 10.1200/JCO.2014.58.1736
The latest version is at http://jco.ascopubs.org/cgi/doi/10.1200/JCO.2014.58.1736

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Impact of Specific Epidermal Growth Factor Receptor (*EGFR*) Mutations and Clinical Characteristics on Outcomes After Treatment With EGFR Tyrosine Kinase Inhibitors Versus Chemotherapy in *EGFR*-Mutant Lung Cancer: A Meta-Analysis

Chee Khoon Lee, Yi-Long Wu, Pei Ni Ding, Sarah J. Lord, Akira Inoue, Caicun Zhou, Tetsuya Mitsudomi, Rafael Rosell, Nick Pavlakis, Matthew Links, Val Gebski, Richard J. Gralla, and James Chih-Hsin Yang

Listen to the podcast by Dr Oxnard at www.jco.org/podcasts

Chee Khoon Lee, Pei Ni Ding, Sarah J. Lord, and Val Gebski, National Health and Medical Research Council Clinical Trials Centre, The University of Sydney; Chee Khoon Lee and Matthew Links, Cancer Care Centre, St George Hospital; Pei Ni Ding, Liverpool Hospital; Sarah J. Lord, School of Medicine, The University of Notre Dame; Nick Pavlakis, Royal North Shore Hospital, Sydney, Australia; Yi-Long Wu, Guangdong Lung Cancer Institute, Guangdong General Hospital and Guangdong Academy of Medical Sciences, Guangdong; Caicun Zhou, Shanghai Pulmonary Hospital, School of Medicine, Tongji University, Shanghai, China; Akira Inoue, Tohoku University Hospital, Sendai; Tetsuya Mitsudomi, Kinki University School of Medicine, Osaka-Sayama, Japan; Rafael Rosell, Catalan Institute of Oncology, Germans Trias i Pujol Health Sciences Institute and Hospital, Barcelona, Spain; Richard J. Gralla, Albert Einstein College of Medicine, Jacobi Medical Center, Bronx, NY; James Chih-Hsin Yang, Graduate Institute of Oncology, National Taiwan University, and National Taiwan University Hospital, Taipei, Taiwan.

Published online ahead of print at www.jco.org on April 20, 2015.

Presented in part at the 15th World Conference on Lung Cancer, Sydney, Australia, October 27-30, 2013.

Authors' disclosures of potential conflicts of interest are found in the article online at www.jco.org. Author contributions are found at the end of this article.

Corresponding author: James Chih-Hsin Yang, MD, PhD, Department of Oncology, National Taiwan University Hospital, National Taiwan University College of Medicine, Taipei 10051, Taiwan; e-mail: chihyang@ntu.edu.tw.

0732-183X/15/3399-1/$20.00

DOI: 10.1200/JCO.2014.58.1736

ABSTRACT

Purpose
We examined the impact of different epidermal growth factor receptor (*EGFR*) mutations and clinical characteristics on progression-free survival (PFS) in patients with advanced *EGFR*-mutated non–small-cell lung cancer treated with EGFR tyrosine kinase inhibitors (TKIs) as first-line therapy.

Patients and Methods
This meta-analysis included randomized trials comparing EGFR TKIs with chemotherapy. We calculated hazard ratios (HRs) and 95% CIs for PFS for the trial population and prespecified subgroups and calculated pooled estimates of treatment efficacy using the fixed-effects inverse-variance-weighted method. All statistical tests were two sided.

Results
In seven eligible trials (1,649 patients), EGFR TKIs, compared with chemotherapy, significantly prolonged PFS overall (HR, 0.37; 95% CI, 0.32 to 0.42) and in all subgroups. For tumors with exon 19 deletions, the benefit was 50% greater (HR, 0.24; 95% CI, 0.20 to 0.29) than for tumors with exon 21 L858R substitution (HR, 0.48; 95% CI, 0.39 to 0.58; $P_{interaction} < .001$). Never-smokers had a 36% greater benefit (HR, 0.32; 95% CI, 0.27 to 0.37) than current or former smokers (HR, 0.50; 95% CI, 0.40 to 0.63; $P_{interaction} < .001$). Women had a 27% greater benefit (HR, 0.33; 95% CI, 0.28 to 0.38) than men (HR, 0.45; 95% CI, 0.36 to 0.55; treatment-sex interaction $P = .02$). Performance status, age, ethnicity, and tumor histology did not significantly predict additional benefit from EGFR TKIs.

Conclusion
Although EGFR TKIs significantly prolonged PFS overall and in all subgroups, compared with chemotherapy, greater benefits were observed in those with exon 19 deletions, never-smokers, and women. These findings should enhance drug development and economic analyses, as well as the design and interpretation of clinical trials.

J Clin Oncol 33. © 2015 by American Society of Clinical Oncology

INTRODUCTION

Advanced non–small-cell lung cancer (NSCLC) with activating mutations in the epidermal growth factor receptor (*EGFR*) gene is a distinct subtype of disease that is characterized by a high tumor response rate when treated with small-molecule EGFR tyrosine kinase inhibitors (TKIs). Randomized trials[1-8] and meta-analyses[9-11] have consistently demonstrated longer progression-free survival (PFS) with EGFR TKI therapy compared with chemotherapy.

Deletions in exon 19 and substitution of leucine for arginine (L858R) in exon 21 of the *EGFR* gene (so-called common mutations) constitute approximately 90% of all *EGFR* mutations that are detected in patients with advanced NSCLC who are enrolled onto randomized trials.[1,2,6,7] Common and uncommon mutation status is used as a stratification factor in many EGFR TKI trials. Although the two common mutations have been regarded as similar in predicting the benefit of EGFR TKIs, subgroup analyses of two studies[6,8] suggested that the benefit of EGFR TKIs is greater in exon 19 deletion

than in exon 21 L858R substitution tumors. However, these findings have not been consistently observed in other trials.[2-5,7]

In the landmark NCIC Clinical Trials Group study BR.21,[12] Asian origin, adenocarcinoma histology, never smoking, and erlotinib were associated with improved overall survival (OS). Subsequent molecular analysis also showed that the benefit of erlotinib was strongly associated with *EGFR* mutation in this trial, and *EGFR* mutations were also more commonly detected in women, patients of Asian origin, patients with adenocarcinoma, and never-smokers.[13,14] Among patients with *EGFR* mutations, the influence of these clinical characteristics on the additional benefit of EGFR TKIs is unknown.

Individual randomized trials have not been designed nor adequately powered to demonstrate a treatment difference between subgroups of patients with these common mutations and other clinicopathologic characteristics. Identifying such factors may be important for future clinical trial design and development of newer generations of EGFR TKIs. To address these questions, this study was designed with the primary objective of testing the hypothesis that the relative effect on PFS of first-line therapy with EGFR TKIs versus chemotherapy is affected by mutation type. Secondary objectives were to test for interactions between clinical characteristics (age, sex, ethnicity, smoking status, performance status, tumor histology) that might be associated with EGFR TKI benefit in a population with *EGFR* mutations.

Ideally, a meta-analysis of randomized trials with OS as the primary end point will address these questions. However, in all of these trials, the effect of EGFR TKIs on OS has been diminished for two reasons: first, nearly all of the patients who were randomly assigned to chemotherapy crossed over to receive EGFR TKIs after disease progression, and second, EGFR TKIs are commercially available outside of clinical trial settings. Furthermore, unlike with EGFR TKIs, the benefit of chemotherapy diminished in second-line as compared with first-line settings. For these reasons, we performed this meta-analysis of PFS outcome using randomized trial data from patients undergoing first-line treatment with first-and second-generation EGFR TKIs.

PATIENTS AND METHODS

Study Eligibility and Identification

Eligible studies were identified from our previous broad systematic review that assessed the effectiveness of EGFR TKIs by *EGFR* mutation status.[9] The included studies were randomized trials that compared EGFR TKIs against platinum-based combination chemotherapy in adult patients with good performance status who did not receive any systemic therapy for their histologically or cytologically confirmed, newly diagnosed advanced NSCLC with sensitizing *EGFR* mutations. In brief, we updated our bibliographic search of MEDLINE, EMBASE, CANCERLIT, and the Cochrane Central Register of Controlled Trials (CENTRAL) databases for articles published in English between January 1, 2004, and February 28, 2014, using the following search terms: lung neoplasms, non–small-cell lung cancer, gefitinib, erlotinib, afatinib, EGFR, meta-analysis, systematic review, randomized, and clinical trials. To identify unpublished studies, we also searched abstracts from conference proceedings of the American Society of Clinical Oncology, the European Society for Medical Oncology, and the World Lung Cancer Conference. Individual study sponsors and study investigators were contacted for conference presentation slides whenever slides were unavailable.

Data Extraction

For each included trial, we extracted the trial name, year of publication or conference presentation, clinicopathologic characteristics, type of chemotherapy, and type of EGFR TKIs. We also retrieved treatment estimates for these subgroups: age (< 65 *v* ≥ 65 years), sex (female *v* male), ethnicity (Asian *v* non-Asian), smoking status (never-smoker *v* current or former smoker), Eastern Cooperative Oncology Group (ECOG) performance status (0 and 1 *v* 2), tumor histology (adenocarcinoma *v* other), and *EGFR* mutation (exon 19 deletion *v* exon 21 L858R substitution) subtype. Data were extracted independently by two authors (P.N.D. and C.K.L.), and discrepancies were resolved by consensus that included a third author (S.J.L.). Risk of bias for PFS analysis in each trial was assessed by examining the methods used in random assignment, allocation concealment, outcome assessments, handling of patient attrition, use of intention-to-treat analysis, and handling of missing data for subgroup analyses.

Statistical Analyses

We extracted the hazard ratios (HRs) and 95% CIs for the overall cohort and subgroups. Data from independent assessment of PFS were used in preference to investigator assessment whenever both types of review were available. We used the fixed-effects inverse-variance-weighted method to pool the results from the studies and to estimate the size of the treatment benefit. Tests for interaction were used to assess differences in treatment effect across subgroups as defined by their baseline clinicopathologic characteristics.

Subgroups with statistically significant heterogeneity in treatment effect were examined further using individual patient data from four trials: NEJ002 (North East Japan 002),[2,15] OPTIMAL,[4] EURTAC (European Tarceva Versus Chemotherapy),[5] and WJTOG (West Japan Thoracic Oncology Group) trial 3405.[3,16] We re-estimated the HRs and 95% CIs in multivariable analyses for the treatment effect for each of these subgroups after adjusting for the other baseline characteristics. We repeated the tests for interaction on the basis of the adjusted HRs to assess differences in treatment effect.

Comparisons between *EGFR* mutations with exon 19 deletions versus exon 21 L858R substitution, with respect to baseline characteristics, involved data from the four trials.[2-5,15,16] The Kaplan-Meier approach was used to examine the difference in PFS between exon 19 deletion and exon 21 L858R substitution in patients who were randomly assigned to the chemotherapy and EGFR TKIs arms separately, and univariable Cox regressions were used to estimate the HRs and 95% CIs.

We performed three sensitivity analyses in which, first, studies were excluded if they reported highly significant subgroup differences in the treatment effect, given that such studies might skew the results if there was selective reporting of chance positive findings; second, the analysis was limited to first-generation EGFR TKIs (gefitinib and erlotinib) because we recognized that there might be differences in efficacy between first- and second-generation EGFR TKIs (afatinib); and third, studies were excluded if the median PFS of the chemotherapy arm differed substantially from that of other included trials because we recognized that there might be differences in efficacy between the different types of platinum combination chemotherapies.

Publication bias was evaluated using the approach of Gleser and Olkin,[17] with an examination of a funnel plot of the effect size for each subgroup of the trial against the reciprocal of its SE.

We used the χ^2 Cochran Q test to detect any heterogeneity across the different studies and between subgroups. The nominal level of significance was set at 5%. All 95% CIs were two sided.

RESULTS

We identified seven eligible studies[2-8,15,18] for inclusion in this meta-analysis (Fig 1). Trial data were obtained from published manuscripts and conference abstracts for three trials.[6-8] Updated individual patient data from the NEJ002[2,15] and OPTIMAL[4] trials were used for subgroup results. Individual patient data with longer follow-up than previously published for EURTAC[5] and WJTOG 3405[3,16] trials were used. Data that were based on independent reviews for PFS were used

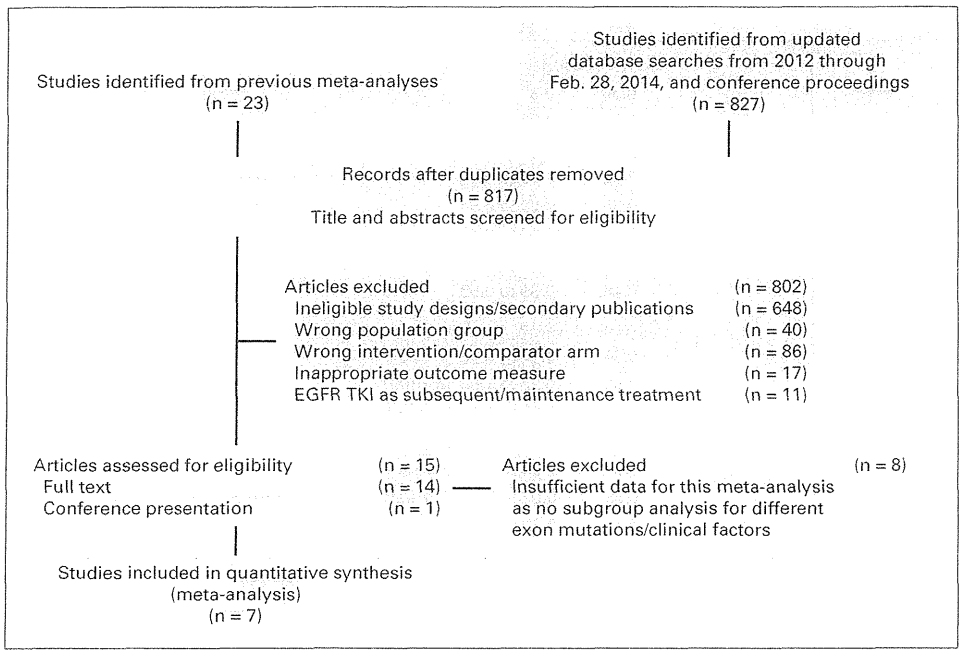


Fig 1. Flow diagram showing inclusion and exclusion of studies. EGFR, epidermal growth factor receptor; TKI, tyrosine kinase inhibitor.

for two studies.[6,7] Hoffmann-La Roche provided unpublished subgroup data for the ENSURE trial that was based on investigator assessment only.[8] All included trials were open label. Risk of bias was assessed as unclear in one unpublished trial,[8] and low for all other studies, although one trial[4] did not include independent review of disease progression.

A total of 1,649 patients participated in these trials. All trials except NEJ002,[2] LUX-Lung 3,[6] and LUX-Lung 6[7] recruited only patients with the two common *EGFR* mutations, exon 19 deletions and exon 21 L858R substitution. Other clinicopathologic characteristics of patients are summarized in Table 1.

Benefit of EGFR TKIs for PFS

Of the 1,649 patients, 950 (58%) had been randomly assigned to EGFR TKIs, and 699 (42%) patients had been randomly assigned to chemotherapy. Treatment with EGFR TKIs compared with chemotherapy was statistically significantly associated with a 63% reduction in the risk of disease progression or death (HR, 0.37; 95% CI, 0.32 to 0.42; $P < .001$).

Subgroup Analyses

Of the 1,558 patients with common mutations, 872 (56%) patients had exon 19 deletions and 686 (44%) had exon 21 L858R

Table 1. Characteristics of Patients in Constituent Trials

Study Name, Year	Treatment Comparison	Median PFS (months)	No. of Patients	Exon 19 Deletion (%)	Exon 21 L858R Substitution (%)	Age < 65 Years (%)	ECOG PS 0 and 1 (%)	Asian (%)	Women (%)	Never-Smoker (%)	Adenocarcinoma (%)
NEJ002, 2010, 2013[2,15*]	Gefitinib *v* CP	10.8 *v* 5.4	224†	51	43	49	99	100	63	62	93
WJTOG 3405, 2010, 2012[3,16]	Gefitinib *v* CisD	9.6 *v* 6.5	172	51	49	53	100	100	69	69	97
OPTIMAL, 2011, 2012[4,18]	Erlotinib *v* CG	13.1 *v* 4.6	154	53	47	75	94	100	59	71	87
EURTAC, 2012[5]	Erlotinib *v* platinum-G or platinum-D	9.7 *v* 5.2	173	66	34	49	86	0	73	69	92
LUX-Lung 3, 2013[6*]	Afatinib *v* CisPem	11.1 *v* 6.9	345	49	40	61	100	72	65	68	100
LUX-Lung 6, 2014[7*]	Afatinib *v* CisG	11.0 *v* 5.6	364	51	38	76	100	100	65	77	100
ENSURE, 2014[8‡]	Erlotinib *v* CisG	11.0 *v* 5.5	217	54	45	79	94	100	61	71	94

Abbreviations: CG, carboplatin-gemcitabine; CisD, cisplatin-docetaxel; CisG, cisplatin-gemcitabine; CisPem, cisplatin-pemetrexed; CP, carboplatin-paclitaxel; ECOG, Eastern Cooperative Oncology Group; EURTAC, European Tarceva Versus Chemotherapy; NEJ002, North East Japan 002; PFS, progression-free survival; PS, performance status; WJTOG, West Japan Thoracic Oncology Group.

*Includes patients with uncommon mutations of the *EGFR* gene.

†NEJ002 recruited a total of 228 patients; PFS outcome was only reported for 224 patients.

‡Reported in abstract only.

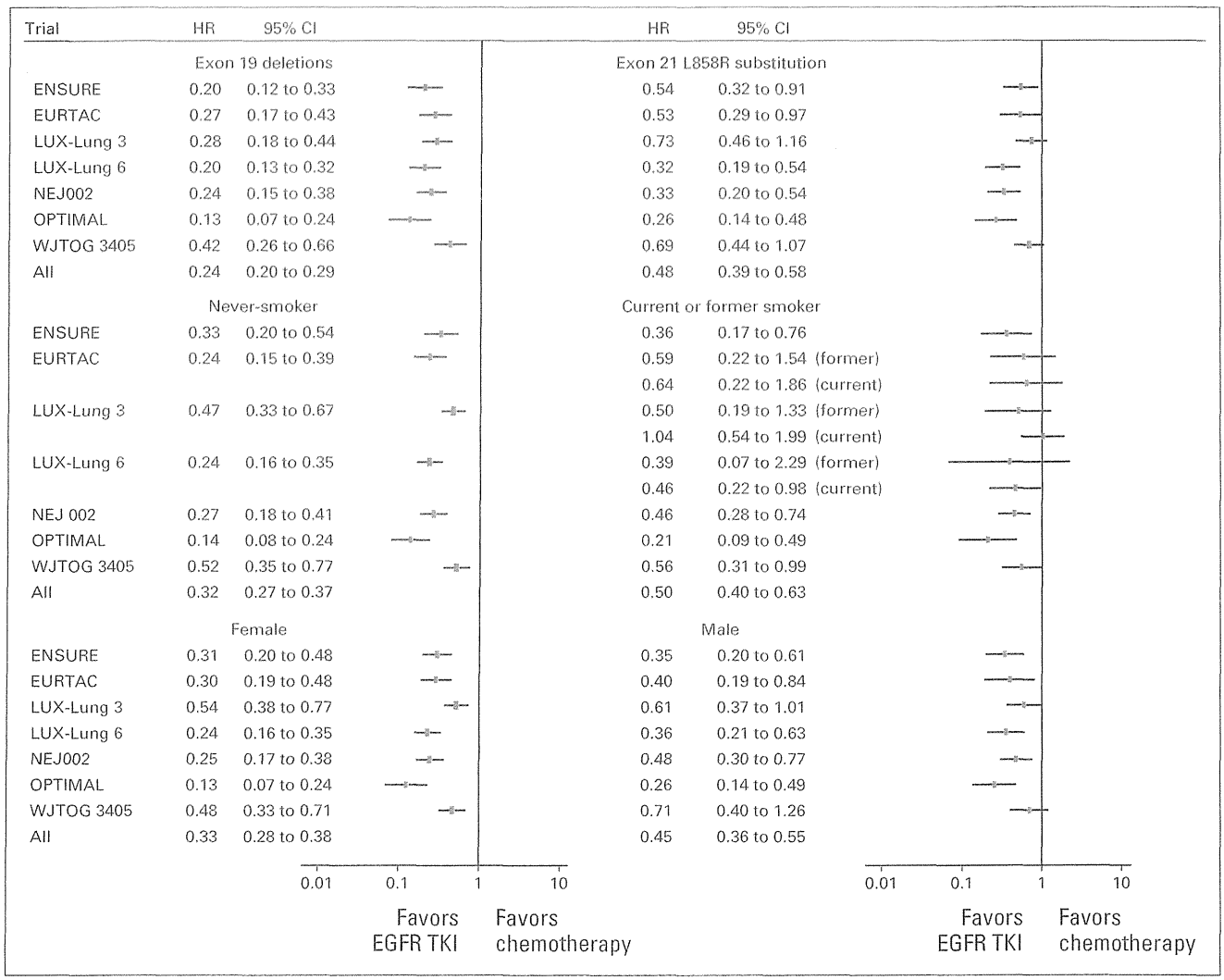


Fig 2. Forest plot of the effect of treatment on progression-free survival in subgroups of patients according to mutations of the epidermal growth factor receptor (*EGFR*) gene, smoking status, and sex. Hazard ratios (HRs) for each trial are represented by the squares, and the horizontal line crossing the square represents the 95% CI. The diamonds represent the estimated overall effect based on the meta-analysis fixed effect. All statistical tests were two sided. EURTAC, European Tarceva Versus Chemotherapy; NEJ002, North East Japan 002; TKI, tyrosine kinase inhibitor; WJTOG, West Japan Thoracic Oncology Group.

substitution. In the subgroup with exon 19 deletions, the pooled HR for PFS was 0.24 (95% CI, 0.20 to 0.29; $P < .001$). In the exon 21 L858R substitution subgroup, the pooled HR for PFS was 0.48 (95% CI, 0.39 to 0.58; $P < .001$). Compared with chemotherapy, treatment with EGFR TKIs demonstrated 50% greater benefit in exon 19 deletions than in exon 21 L858R substitution (interaction $P < .001$; Fig 2).

Of the 1,649 patients, most were never-smokers (n = 1,155; 70%) and 494 (30%) were current or former smokers. Among the never-smokers, the pooled HR for PFS was 0.32 (95% CI, 0.27 to 0.37; $P < .001$). Among the current or former smokers, the pooled HR for PFS was 0.50 (95% CI, 0.40 to 0.63; $P < .001$). Compared with chemotherapy, treatment with EGFR TKIs demonstrated a 36% greater benefit in never-smokers than current or former smokers (interaction $P = .002$; Fig 2).

Most patients (n = 1,073; 65%) were women; 576 (35%) were men. Among the women, the pooled HR for PFS was 0.33 (95% CI, 0.28 to 0.38; $P < .001$). Among the men, the pooled HR for PFS was 0.45 (95% CI, 0.36 to 0.55; $P < .001$). Compared with chemotherapy, EGFR TKI treatment demonstrated a 27% greater benefit in women than men (interaction $P = .02$; Fig 2).

In multivariable analysis using data from the four trials,[2-5,15,16] the pooled HRs for PFS were 0.26 and 0.44, adjusted for smoking status and sex, for exon 19 deletions and exon 21 L858R substitution subgroups, respectively (interaction $P = .004$). There was negligible difference in the result between unadjusted and adjusted HRs (exon 19 deletions: unadjusted pooled HR, 0.26; exon 21 L858R substitution: unadjusted pooled HR, 0.45; interaction $P = .004$). Table 2 compares the unadjusted and adjusted HRs of treatment effect to assess any potential inter-related impact of type of *EGFR* mutation, sex, and smoking on benefit with EGFR TKIs.

The improvement in PFS with EGFR TKI treatment compared with chemotherapy did not differ by ethnicity (interaction $P = .37$), age (interaction $P = .27$), tumor histologic subtype (interaction $P = .59$), or performance status (interaction $P = .85$; Fig 3).

Table 2. Unadjusted and Adjusted Treatment Effect of EGFR TKIs Versus Chemotherapy in Four Clinical Trials

Subgroup	Unadjusted Analysis HR	Unadjusted Analysis 95% CI	Adjusted Analysis HR	Adjusted Analysis 95% CI
Exon 19 deletions				
EURTAC	0.27	0.17 to 0.43	0.25*	0.15 to 0.41
NEJ002	0.24	0.15 to 0.38	0.24*	0.15 to 0.38
OPTIMAL	0.13	0.07 to 0.25	0.12*	0.06 to 0.22
WJTOG 3405	0.42	0.26 to 0.68	0.46*	0.28 to 0.76
Pooled result	0.26	0.20 to 0.34	0.26	0.20 to 0.33
Exon 21 L858R substitution				
EURTAC	0.53	0.29 to 0.97	0.51*	0.28 to 0.94
NEJ002	0.33	0.20 to 0.54	0.33*	0.20 to 0.55
OPTIMAL	0.26	0.14 to 0.49	0.23*	0.12 to 0.45
WJTOG 3405	0.69	0.44 to 1.07	0.69*	0.44 to 1.08
Pooled result	0.45	0.34 to 0.58	0.44	0.34 to 0.58
Treatment-*EGFR* mutation interaction	$P = .004$		$P = .004$	
Never-smoker				
EURTAC	0.24	0.15 to 0.39	0.23†	0.14 to 0.38
NEJ002	0.27	0.18 to 0.41	0.24†	0.16 to 0.37
OPTIMAL	0.14	0.08 to 0.25	0.14†	0.08 to 0.25
WJTOG 3405	0.52	0.35 to 0.77	0.52†	0.34 to 0.79
Pooled result	0.29	0.24 to 0.37	0.28	0.22 to 0.35
Current or former smoker				
EURTAC (former)	0.59	0.22 to 1.54	0.67†	0.25 to 1.78
EURTAC (current)	0.64	0.22 to 1.86	0.56†	0.19 to 1.71
NEJ002	0.46	0.28 to 0.74	0.45†	0.28 to 0.73
OPTIMAL	0.21	0.09 to 0.49	0.20†	0.08 to 0.47
WJTOG 3405	0.56	0.31 to 0.99	0.57†	0.32 to 1.02
Pooled result	0.46	0.34 to 0.62	0.46†	0.34 to 0.62
Treatment-smoking interaction	$P = .02$		$P = .01$	
Women				
EURTAC	0.30	0.19 to 0.48	0.29‡	0.18 to 0.47
NEJ002	0.25	0.17 to 0.38	0.21‡	0.14 to 0.33
OPTIMAL	0.13	0.07 to 0.24	0.13‡	0.07 to 0.24
WJTOG 3405	0.48	0.33 to 0.71	0.50‡	0.33 to 0.76
Pooled result	0.30	0.24 to 0.38	0.28	0.22 to 0.36
Men				
EURTAC	0.40	0.19 to 0.84	0.37‡	0.17 to 0.81
NEJ002	0.48	0.30 to 0.77	0.45‡	0.28 to 0.74
OPTIMAL	0.26	0.14 to 0.50	0.23‡	0.12 to 0.45
WJTOG 3405	0.71	0.40 to 1.26	0.69‡	0.39 to 1.22
Pooled result	0.46	0.34 to 0.61	0.43	0.32 to 0.58
Treatment-sex interaction	$P = .02$		$P = .03$	

Abbreviations: EGFR, epidermal growth factor receptor; EURTAC, European Tarceva Versus Chemotherapy; HR, hazard ratio; NEJ002, North East Japan 002; TKI, tyrosine kinase inhibitor; WJTOG, West Japan Thoracic Oncology Group.
*HR (EGFR TKI *v* chemotherapy) adjusted for smoking status and sex.
†HR (EGFR TKI *v* chemotherapy) adjusted for sex and type of *EGFR* mutation.
‡HR (EGFR TKI *v* chemotherapy) adjusted for smoking status and type of *EGFR* mutation.

Benefit of EGFR TKIs for OS

At the point of data cutoff for this analysis, several trials had reported preliminary OS data and had patients still in active follow-up. The data for OS remained immature for many of these studies. The OS data for the ENSURE trial was unavailable.[8] With the available preliminary OS data from the remaining six trials, treatment with EGFR TKIs compared with chemotherapy was not statistically significantly associated with reduction in the risk of death (HR, 1.01; 95% CI, 0.86 to 1.19; $P = .88$).

Association Between Mutations and Baseline Clinical Characteristics

In four trials,[2-5,15,16] there were no significant correlations between *EGFR* mutation type and age, performance status, sex, histology, or smoking status (Table 3).

Prognostic Outcomes for Patients With Common Mutations

Of the 348 patients in the four trials[2-5,15,16] who were randomly assigned to chemotherapy, those with exon 21 L858R substitution (n = 158) had a median PFS of 6.1 months, which was statistically significantly longer than those with exon 19 deletions (n = 190), who had a median PFS of 5.1 months (HR, 0.70; 95% CI, 0.56 to 0.89; $P = .003$). In comparison, of the 362 patients who were randomly assigned to EGFR TKIs in these trials, patients with exon 21 L858R substitution (n = 154) had a median PFS of 10.0 months, which was statistically significantly shorter than that of patients with exon 19 deletions (n = 208), who had a median PFS of 11.8 months (HR, 1.39; 95% CI, 1.10 to 1.76; $P = .006$).

Publication Bias

A funnel plot of the effect size for each subgroup category of the trial against the precision showed no asymmetry (not shown). A formal test[17] for potential publication bias yielded no potential unpublished studies.

Sensitivity Analyses

Two trials[6,8] individually demonstrated greater PFS benefit for EGFR TKIs versus chemotherapy in tumors with exon 19 deletions compared with those with exon 21 L858R substitution; therefore, we excluded these studies and observed consistent results (HR, 0.24 *v* 0.42; interaction $P < .001$; Appendix Fig A1, online only).

Restricting our analyses to trials of first-generation reversible EGFR TKIs, erlotinib[4,5,8,18] and gefitinib[2,3,15,16] (Appendix Fig A2, online only), we also found consistent results: greater benefit with EGFR TKIs for exon 19 deletions (interaction $P < .001$), never-smokers (interaction $P = .03$), and women (interaction $P = .03$).

Two trials[3,6] individually demonstrated median PFS greater than 6 months in the chemotherapy arm. Given that this was a longer PFS than reported in other studies (Table 1), we excluded these two studies and observed consistent results: greater benefit for EGFR TKIs for exon 19 deletions (interaction $P < .001$), never-smokers (interaction $P = .003$), and women (interaction $P = .01$; Appendix Fig A3, online only).

DISCUSSION

Treatment with EGFR TKIs compared with chemotherapy is associated with a 63% overall reduction in the risk of disease progression or death. Furthermore, the relative effect of EGFR TKIs compared with chemotherapy on PFS is 50% greater for patients with exon 19 deletions than for those with exon 21 L858R substitution. Other crucial findings include a 36% greater PFS benefit for never-smokers than

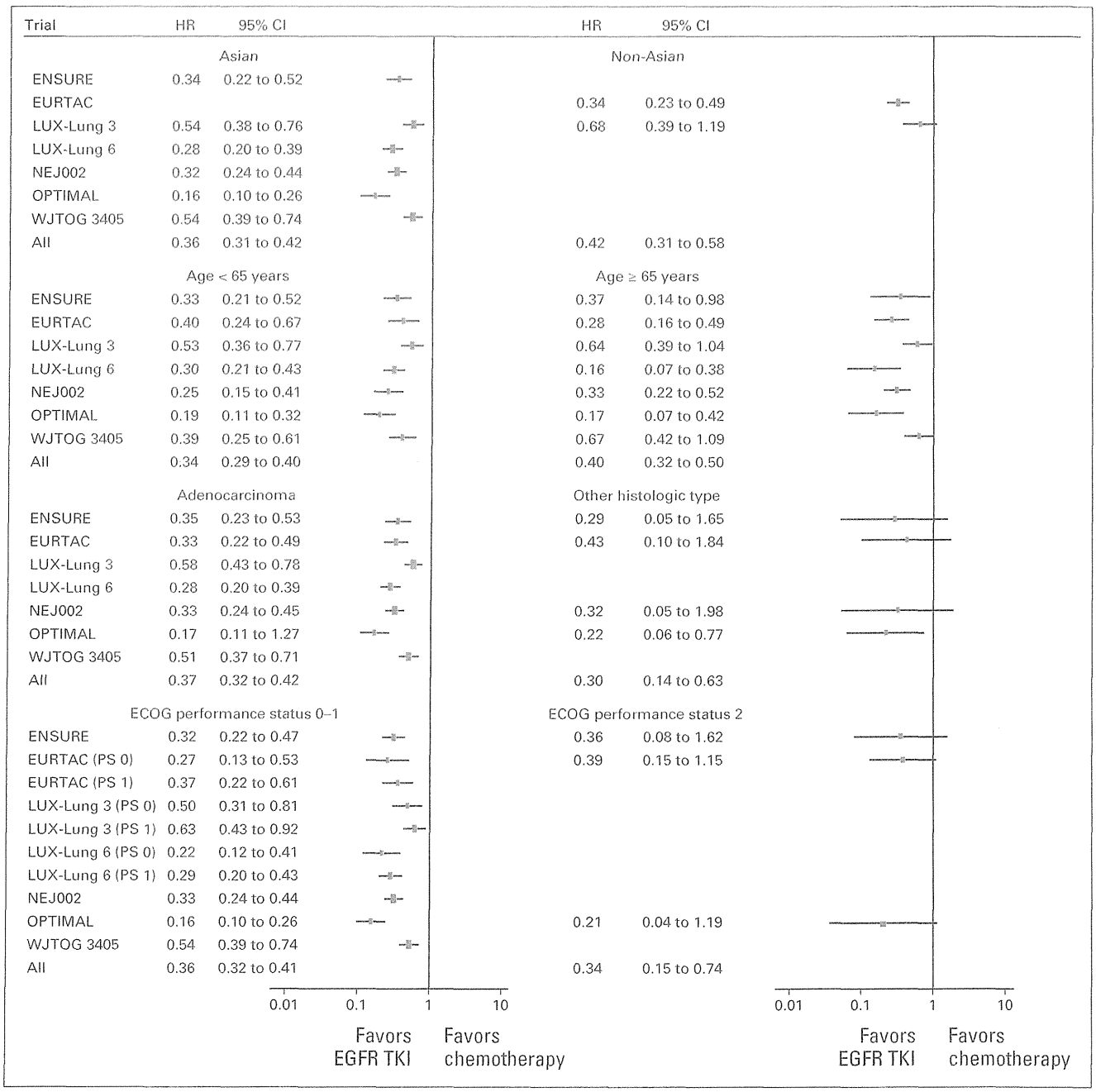


Fig 3. Forest plot of the effect of treatment on progression-free survival in subgroups of patients according to ethnicity, age, tumor histologic subtype, and performance status (PS). Hazard ratios (HRs) for each trial are represented by the squares, and the horizontal line crossing the square represents the 95% CI. The diamonds represent the estimated overall effect based on the meta-analysis fixed effect. All statistical tests were two sided. ECOG, Eastern Cooperative Oncology Group; EGFR, epidermal growth factor receptor; EURTAC, European Tarceva Versus Chemotherapy; NEJ002, North East Japan 002; TKI, tyrosine kinase inhibitor; WJTOG, West Japan Thoracic Oncology Group.

current or former smokers and a 27% greater PFS benefit for women than men with EGFR TKIs compared with chemotherapy.

Consistent with previous studies, patients with exon 19 deletions have a longer OS than those with exon 21 L858R substitution after gefitinib or erlotinib therapy.[19,20] In contrast, in patients who are not treated with EGFR TKIs, exon 21 L858R substitution, rather than exon 19 deletions, has been associated with longer OS.[14] Using data from four trials,[2-5,15,16] we found that patients randomly assigned to chemotherapy who had exon 21 L858R substitution had statistically significantly longer PFS than those with exon 19 deletions (median PFS, 6.1 *v* 5.1 months; $P = .003$). This indicates that patients who harbor exon 19 deletions and are not treated with EGFR TKIs have a poorer prognosis than those with exon 21 L858R substitution. Treatment with EGFR TKIs improves the prognosis more in those with exon 19 deletions than in those with exon 21 L858R substitution (median PFS, 11.8 *v* 10.0 months; $P = .006$).

Table 3. Association Between Baseline Characteristics and Exon 19 Deletion or Exon 21 L858R Substitution: Pooled Data From Four Clinical Trials

	Exon 19 Deletion (n = 401)		Exon 21 L858R Substitution (n = 313)		
Characteristic	No.	%	No.	%	*P*
Age, years					.20
< 65	233	58	166	53	
≥ 65	168	42	147	47	
ECOG PS					.32
0	186	46	136	44	
1	191	48	164	52	
2	24	6	13	4	
Sex					.81
Female	268	67	206	66	
Male	133	33	107	34	
Smoking					.81
Never	268	67	212	68	
Ever	133	33	101	32	
Histologic subtype					.11
Adenocarcinoma	377	94	284	91	
Other	24	6	29	9	

Abbreviations: EGOG, Eastern Cooperative Oncology Group; PS, performance status.

The associations between different *EGFR* mutations and baseline clinicopathologic characteristics remain unclear. Several studies report that exon 21 L858R substitution is more frequently associated with female sex, never smoking, and having adenocarcinoma.[21-23] Use of the largest pooled individual patient data set of common mutations (n = 714) from four trials[2-5,15,16] failed to detect any association between the type of mutation and smoking status ($P = .81$), histology ($P = .11$), or sex ($P = .81$).

Our finding that smoking status modifies EGFR TKI benefit is also supported by existing studies. Smoking was found to be independently associated with poorer tumor response with gefitinib.[24] Smoking was also associated with significantly less drug exposure after ingestion of erlotinib.[25] A phase I study[26] of smokers reported a maximum tolerated erlotinib dose of 300 mg, which was much higher than the dose of 150 mg per day used in randomized trials.[4,5,8] Whether this metabolic difference is the true reason for the PFS difference or whether other factors are involved has yet to be determined, and further research is warranted.

Another interesting finding was that women had a 27% greater PFS benefit with EGFR TKIs than men. The benefit of EGFR TKIs in women has been previously attributed to the higher rate of *EGFR* mutations in women.[14] In this meta-analysis involving only trials conducted in populations with *EGFR* activating mutations, a difference in PFS benefit on the basis of sex was still detected. As a majority of the nonsmokers were also women in these trials, it is possible that smoking is confounding the interaction between sex and EGFR TKI efficacy. However, multivariable analysis performed using individual patient data from four trials[2-5,15,16] suggests that the predictive effect of sex is largely independent of smoking status and *EGFR* mutation type (Table 2). We acknowledge that there may be a difference between current and former smokers, but our analysis does not discriminate between these two cohorts of patients.

This meta-analysis has several strengths. We performed a comprehensive review, used the most up-to-date published data, and contacted individual investigators or trial sponsors to obtain relevant unpublished data. Another strength is that individual patient data from four trials[2-5,15,16] were available to investigate the relationships between different *EGFR* mutations and baseline clinical characteristics, for multivariable adjustment, and for prognostic analyses.

There are also limitations of this study. We have not reported the treatment effects within subgroups for OS because many of the trials have yet to report mature OS data. In a recently presented pooled analysis of two randomized trials, OS was longer with afatinib than chemotherapy, and a statistically significant prolongation of OS was reported in tumors with exon 19 deletions but not exon 21 L858R substitution.[27] It remains unknown whether there would be a similar finding in first-generation EGFR TKI trials. We restricted our study to common *EGFR* mutations, and the predictive value of uncommon mutations remains unknown. We are currently planning an individual patient data meta-analysis using all randomized trials with mature OS data to address the limitations of our current work.

Our results have several important clinical and research implications. Our findings will be useful for counseling patients. Our meta-analysis demonstrates that exon 19 deletion and exon 21 L858R substitution mutations have different prognostic and predictive roles and are hence important as a stratification factor in future clinical trials. Further drug development of EGFR TKIs to enhance antitumor activity, particularly for tumors with exon 21 L858R substitution, remains important.

Another potential use of these findings is in economic analyses. With differences in PFS benefits for various subgroups, there will be differences in the costs required to achieve these benefits. In addition, economic factors related to patient screening may also identify greater cost-benefit for different identifiable subgroups.

In conclusion, EGFR TKIs significantly prolong PFS in all patients with advanced NSCLC with *EGFR* mutations compared with chemotherapy. The relative benefits of EGFR TKIs compared with chemotherapy were greatest in patients with exon 19 deletions. Greater PFS benefit with EGFR TKIs compared with chemotherapy was also seen in never-smokers and women. These findings have important implications for clinical trial design and interpretation, economic analyses, and future drug development for *EGFR*-mutated, advanced NSCLC.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at www.jco.org.

AUTHOR CONTRIBUTIONS

Conception and design: Chee Khoon Lee, Yi-Long Wu, Pei Ni Ding, Sarah J. Lord, Akira Inoue, Tetsuya Mitsudomi, Nick Pavlakis, Matthew Links, Val Gebski, Richard J. Gralla, James Chih-Hsin Yang
Collection and assembly of data: Chee Khoon Lee, Yi-Long Wu, Pei Ni Ding, Sarah J. Lord, Akira Inoue, Caicun Zhou, Tetsuya Mitsudomi, Rafael Rosell, James Chih-Hsin Yang
Data analysis and interpretation: Chee Khoon Lee, Yi-Long Wu, Pei Ni Ding, Sarah J. Lord, Akira Inoue, Tetsuya Mitsudomi, Rafael Rosell, Nick

Pavlakis, Matthew Links, Val Gebski, Richard J. Gralla, James Chih-Hsin Yang

Manuscript writing: All authors
Final approval of manuscript: All authors

REFERENCES

1. Mok TS, Wu Y-L, Thongprasert S, et al: Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med 361:947-957, 2009
2. 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
3. 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
4. Zhou C, Wu Y-L, Chen G, et al: Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): A multicentre, open-label, randomised, phase 3 study. Lancet Oncol 12:735-742, 2011
5. Rosell R, Carcereny E, Gervais R, et al: Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): A multicentre, open-label, randomised phase 3 trial. Lancet Oncol 13:239-246, 2012
6. Sequist LV, Yang JC, Yamamoto N, et al: Phase III study of afatinib or cisplatin plus pemetrexed in patients with metastatic lung adenocarcinoma with EGFR mutations. J Clin Oncol 31:3327-3334, 2013
7. Wu Y-L, Zhou C, Hu C-P, et al: Afatinib versus cisplatin plus gemcitabine for first-line treatment of Asian patients with advanced non-small-cell lung cancer harbouring EGFR mutations (LUX-Lung 6): An open-label, randomised phase 3 trial. Lancet Oncol 15:213-222, 2014
8. Wu Y-L, Liam C-K, Zhou C, et al: First-line erlotinib versus cisplatin/gemcitabine (GP) in patients with advanced EGFR mutation-positive non-small-cell lung cancer (NSCLC): Interim analyses from the phase 3, open-label, ENSURE study. J Thorac Oncol 8:S603, 2013
9. Lee CK, Brown C, Gralla RJ, et al: Impact of EGFR inhibitor in non-small cell lung cancer on progression-free and overall survival: A meta-analysis. J Natl Cancer Inst 105:595-605, 2013
10. Bria E, Milella M, Cuppone F, et al: Outcome of advanced NSCLC patients harboring sensitizing EGFR mutations randomized to EGFR tyrosine kinase inhibitors or chemotherapy as first-line treatment: A meta-analysis. Ann Oncol 22:2277-2285, 2011
11. Gao G, Ren S, Li A, et al: Epidermal growth factor receptor-tyrosine kinase inhibitor therapy is effective as first-line treatment of advanced non-small-cell lung cancer with mutated EGFR: A meta-analysis from six phase III randomized controlled trials. Int J Cancer 131:E822-E829, 2012
12. 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
13. Zhu CQ, da Cunha Santos G, Ding K, et al: Role of KRAS and EGFR as biomarkers of response to erlotinib in National Cancer Institute of Canada Clinical Trials Group study BR.21. J Clin Oncol 26:4268-4275, 2008
14. Shigematsu H, Lin L, Takahashi T, et al: Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst 97:339-346, 2005
15. Inoue A, Kobayashi K, Maemondo M, et al: Updated overall survival results from a randomized phase III trial comparing gefitinib with carboplatin-paclitaxel for chemo-naïve non-small cell lung cancer with sensitive EGFR gene mutations (NEJ002). Ann Oncol 24:54-59, 2013
16. Mitsudomi T, Morita S, Yatabe Y, et al: Updated overall survival results of WJTOG 3405, a randomized phase III trial comparing gefitinib (G) with cisplatin plus docetaxel (CD) as the first-line treatment for patients with non-small cell lung cancer harboring mutations of the epidermal growth factor receptor (EGFR). J Clin Oncol 30:485s, 2012 (suppl; abstr 7521)
17. Gleser LJ, Olkin I: Models for estimating the number of unpublished studies. Stat Med 15:2493-2507, 1996
18. Zhou C, Wu YL, Liu X, et al: Overall survival (OS) results from OPTIMAL (CTONG0802), a phase III trial of erlotinib (E) versus carboplatin plus gemcitabine (GC) as first-line treatment for Chinese patients with EGFR mutation-positive advanced non-small cell lung cancer (NSCLC). J Clin Oncol 30:485s, 2012 (suppl; abstr 7520)
19. Riely GJ, Pao W, Pham D, et al: Clinical course of patients with non–small cell lung cancer and epidermal growth factor receptor exon 19 and exon 21 mutations treated with gefitinib or erlotinib. Clin Cancer Res 12:839-844, 2006
20. Jackman DM, Yeap BY, Sequist LV, et al: Exon 19 deletion mutations of epidermal growth factor receptor are associated with prolonged survival in non–small cell lung cancer patients treated with gefitinib or erlotinib. Clin Cancer Res 12:3908-3914, 2006
21. Moriguchi H, Kim T-Y, Sato C: Gefitinib for refractory advanced non-small-cell lung cancer. Lancet 367:299-300, 2006
22. Sugio K, Uramoto H, Ono K, et al: Mutations within the tyrosine kinase domain of EGFR gene specifically occur in lung adenocarcinoma patients with a low exposure of tobacco smoking. Br J Cancer 94:896-903, 2006
23. Sasaki H, Endo K, Takada M, et al: L858R EGFR mutation status correlated with clinicopathological features of Japanese lung cancer. Lung Cancer 54:103-108, 2006
24. Lim ST, Wong EH, Chuah KL, et al: Gefitinib is more effective in never-smokers with non-small-cell lung cancer: Experience among Asian patients. Br J Cancer 93:23-28, 2005
25. Hamilton M, Wolf JL, Rusk J, et al: Effects of smoking on the pharmacokinetics of erlotinib. Clin Cancer Res 12:2166-2171, 2006
26. Hughes AN, O'Brien ME, Petty WJ, et al: Overcoming CYP1A1/1A2 mediated induction of metabolism by escalating erlotinib dose in current smokers. J Clin Oncol 27:1220-1226, 2009
27. Yang JC, Wu YL, Schuler M, et al: Afatinib versus cisplatin-based chemotherapy for EGFR mutation-positive lung adenocarcinoma (LUX-Lung 3 and LUX-Lung 6): Analysis of overall survival data from two randomised, phase 3 trials. Lancet Oncol 16:141-151, 2015

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Impact of Specific Epidermal Growth Factor Receptor (*EGFR*) Mutations and Clinical Characteristics on Outcomes After Treatment With EGFR Tyrosine Kinase Inhibitors Versus Chemotherapy in *EGFR*-Mutant Lung Cancer: A Meta-Analysis

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or jco.ascopubs.org/site/ifc.

Chee Khoon Lee
Research Funding: GlaxoSmithKline (Inst)
Travel, Accommodations, Expenses: Boehringer Ingelheim

Yi-Long Wu
Honoraria: AstraZeneca, Roche, Eli Lilly

Pei Ni Ding
No relationship to disclose

Sarah J. Lord
No relationship to disclose

Akira Inoue
Honoraria: AstraZeneca, Chugai Pharma, Boehringer Ingelheim
Consulting or Advisory Role: AstraZeneca, Boehringer Ingelheim
Research Funding: AstraZeneca, Chugai Pharma, Boehringer Ingelheim

Caicun Zhou
No relationship to disclose

Tetsuya Mitsudomi
Honoraria: AstraZeneca, Chugai, Boehringer Ingelheim, Pfizer, Taiho, Eli Lilly, Daiichi Sankyo
Consulting or Advisory Role: AstraZeneca, Norvatis, Chugai, Boehringer Ingelheim, Pfizer, Roche, Synta, Clovis, Merck Sharpe & Dohme
Research Funding: AstraZeneca (Inst), Chugai (Inst), Boehringer Ingelheim (Inst), Pfizer (Inst), Taiho (Inst), Ono (Inst), Daiichi Sankyo (Inst), Eli Lilly (Inst)

Rafael Rosell
No relationship to disclose

Nick Pavlakis
Honoraria: AstraZeneca, Roche, Boehringer Ingelheim
Consulting or Advisory Role: AstraZeneca, Roche, Boehringer Ingelheim, Bristol-Myers Squibb, Pfizer

Matthew Links
Travel, Accommodations, Expenses: GlaxoSmithKline

Val Gebski
No relationship to disclose

Richard J. Gralla
Consulting or Advisory Role: Eli Lilly, Merck, Pierre Fabre

James Chih-Hsin Yang
Honoraria: Boehringer Ingelheim, Pfizer, Roche/Genentech, AstraZeneca
Consulting or Advisory Role: AstraZeneca, Boehringer Ingelheim (Inst), Roche/Genentech, Merck Serono, Novartis, Bayer, Takeda (Inst), Clovis Oncology
Research Funding: Boehringer Ingelheim (Inst)

Acknowledgment

We thank Hoffmann-La Roche for providing us with unpublished data for this meta-analysis. We acknowledge the editorial support provided by Rhana Pike (National Health and Medical Research Council Clinical Trials Centre).

Appendix

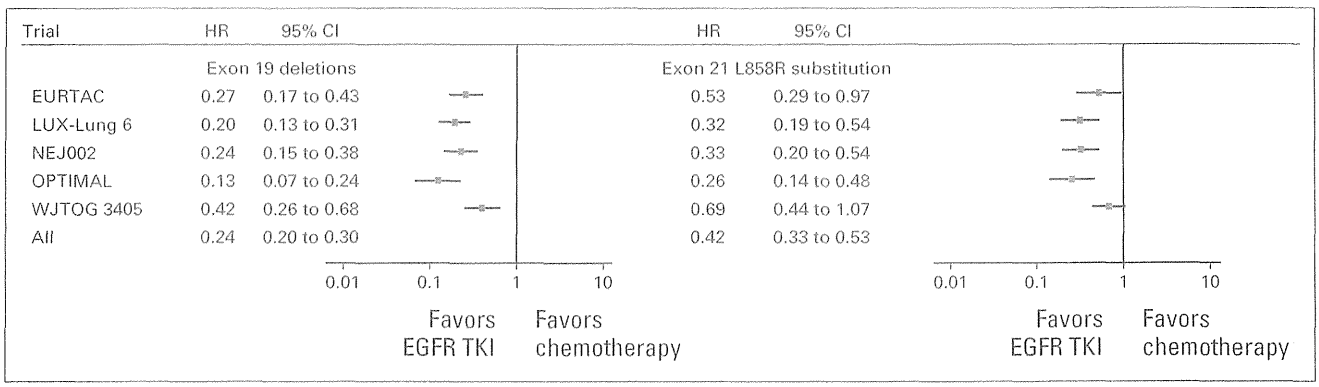


Fig A1. Forest plot of effect of treatment on progression-free survival in subgroups of patients according to different mutations of the epidermal growth factor receptor (*EGFR*), with exclusion of the LUX-Lung 3 and ENSURE trials. Hazard ratios (HRs) for each trial are represented by the squares, and the horizontal line crossing the square represents the 95% CI. The diamonds represent the estimated overall effect based on the meta-analysis fixed effect (all $P < .001$). All statistical tests were two sided. EURTAC, European Tarceva Versus Chemotherapy; NEJ002, North East Japan 002; TKI, tyrosine kinase inhibitor; WJTOG, West Japan Thoracic Oncology Group.

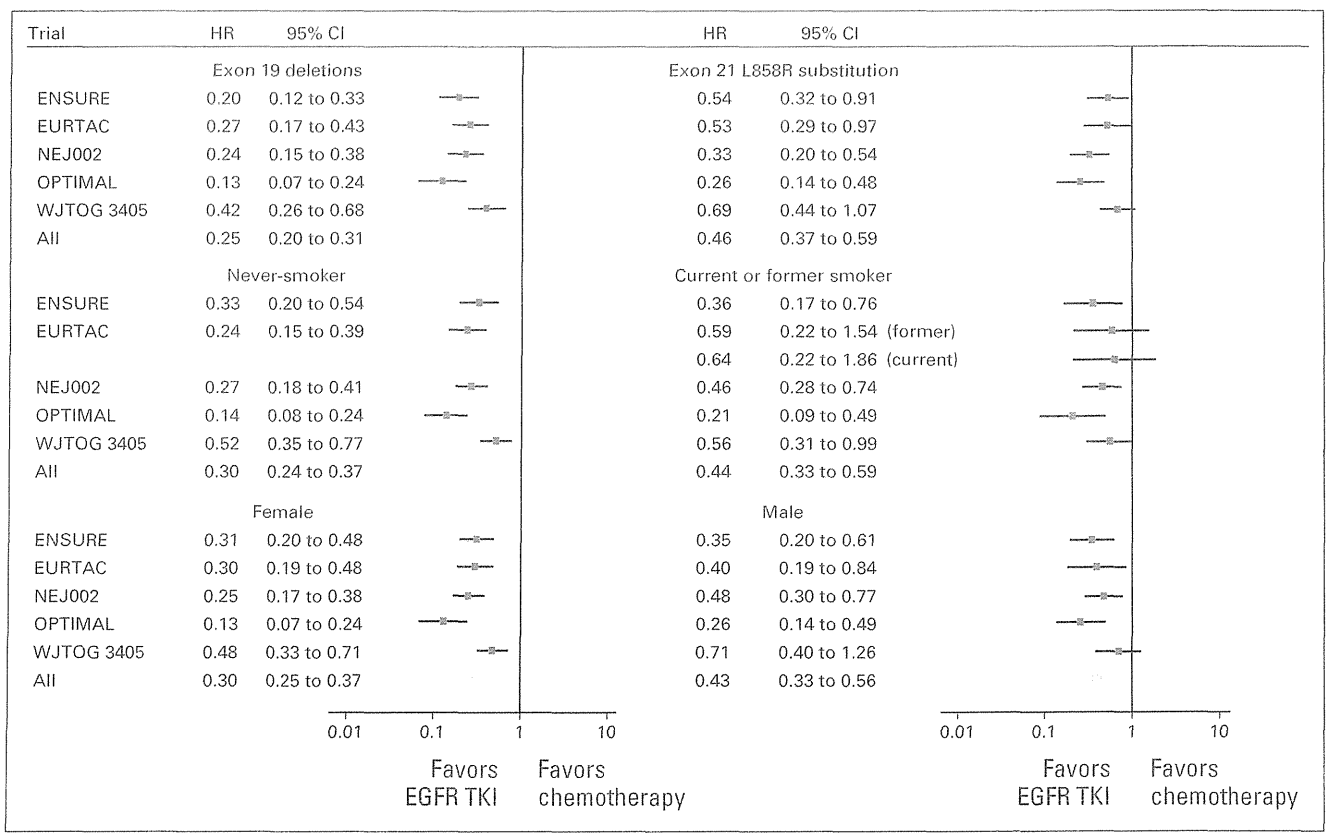


Fig A2. Forest plot of effect of treatment on progression-free survival in subgroups of patients according to mutations of the epidermal growth factor receptor (*EGFR*) gene, smoking status, and sex in gefitinib and erlotinib trials only. Hazard ratios (HRs) for each trial are represented by the squares, and the horizontal line crossing the square represents the 95% CI. The diamonds represent the estimated overall effect based on the meta-analysis of fixed effect (all $P < .001$). All statistical tests were two sided. EURTAC, European Tarceva Versus Chemotherapy; NEJ002, North East Japan 002; TKI, tyrosine kinase inhibitor; WJTOG, West Japan Thoracic Oncology Group.