potential peripheral blood biomarkers in idiopathic pulmonary fibrosis. *PLoS Med* 2008;5:e93.

161. Gilani SR, Vuga LJ, Lindell KO, Gibson KF, Xue J, Kaminski N, Valentine VG, Lindsay EK, George MP, Steele C, *et al.* CD28 down-regulation on circulating CD4 T-cells is associated with poor prognoses of patients with idiopathic pulmonary fibrosis. *PLoS One* 2010;5:e8959.
162. Richards TJ, Kaminski N, Baribaud F, Flavin S, Brodmerkel C, Horowitz D, Li K, Choi J, Vuga LJ, Lindell KO, Klesen M, Zhang Y, Gibson KF. Peripheral blood proteins predict mortality in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2012;185:67–76.
163. McKeown S, Richter AG, O'Kane C, McAuley DF, Thickett DR. MMP expression and abnormal lung permeability are important determinants of outcome in IPF. *Eur Respir J* 2009;33:77–84.
164. Shinoda H, Tasaka S, Fujishima S, Yamasawa W, Miyamoto K, Nakano Y, Kamata H, Hasegawa N, Ishizaka A. Elevated CC chemokine level in bronchoalveolar lavage fluid is predictive of a poor outcome of idiopathic pulmonary fibrosis. *Respiration* 2009;78:285–292.
165. Richter AG, McKeown S, Rathinam S, Harper L, Rajesh P, McAuley DF, Heljasvaara R, Thickett DR. Soluble endostatin is a novel inhibitor of epithelial repair in idiopathic pulmonary fibrosis. *Thorax* 2009;64:156–161.

CHEST Original Research
DIFFUSE LUNG DISEASE

St. George's Respiratory Questionnaire Has Longitudinal Construct Validity in Lymphangioleiomyomatosis

Jeffrey J. Swigris, DO, MS; Hye-Seung Lee, PhD; Marsha Cohen, MD; Yoshikazu Inoue, MD; Joel Moss, MD, PhD, FCCP; Lianne G. Singer, MD, FCCP; Lisa R. Young, MD; and Francis X. McCormack, MD, FCCP; for the National Institutes of Health Rare Lung Diseases Consortium and the Multicenter International Lymphangioleiomyomatosis Efficacy and Safety of Sirolimus Trial Group

Background: **Lymphangioleiomyomatosis (LAM) is an uncommon, progressive, cystic lung disease that causes shortness of breath, hypoxemia, and impaired health-related quality of life (HRQL). Whether St. George's Respiratory Questionnaire (SGRQ), a respiratory-specific HRQL instrument, captures longitudinal changes in HRQL in patients with LAM is unknown.**
Methods: **Using data from the Multicenter International Lymphangioleiomyomatosis Efficacy and Safety of Sirolimus trial, we performed analyses to examine associations between SGRQ scores and values for four external measures (anchors). Anchors included (1) FEV_1, (2) diffusing capacity of the lung for carbon monoxide, (3) distance walked during the 6-min walk test, and (4) serum vascular endothelial growth factor-D.**
Results: **SGRQ scores correlated with the majority of anchor values at baseline, 6 months, and 12 months. Results from longitudinal analyses demonstrated that SGRQ change scores tracked changes over time in values for each of the four anchors. At 12 months, subjects with the greatest improvement from baseline in FEV_1 experienced the greatest improvement in SGRQ scores (Symptoms domain, -13.4 ± 14.6 points; Activity domain, -6.46 ± 8.20 points; Impacts domain, -6.25 ± 12.8 points; SGRQ total, -7.53 ± 10.0 points). Plots of cumulative distribution functions further supported the longitudinal validity of the SGRQ in LAM.**
Conclusions: **In LAM, SGRQ scores are associated with variables used to assess LAM severity. The SGRQ is sensitive to change in LAM severity, particularly when change is defined by FEV_1, perhaps the most clinically relevant and prognostically important variable in LAM. The constellation of results here supports the validity of the SGRQ as capable of assessing longitudinal change in HRQL in LAM.** ***CHEST 2013; 143(6):1671–1678***

Abbreviations: 6MWD = distance walked during 6-min walk test; CDF = cumulative distribution function; DLCO = diffusing capacity of the lung for carbon monoxide; LAM = lymphangioleiomyomatosis; SGRQ = St. George's Respiratory Questionnaire; VEGF-D = serum vascular endothelial growth factor D level

Lymphangioleiomyomatosis (LAM) is an uncommon, progressive lung disease that affects women and occurs either sporadically or in association with tuberous sclerosis complex.[1-3] In LAM, hallmark cystic destruction of the pulmonary parenchyma impairs lung function and induces debilitating dyspnea.[4,5] In most patients, hypoxemia develops within a decade of symptom onset.[6] Given the breathlessness, functional limitation, and need for supplemental oxygen that ultimately develop, it is not surprising that patients with LAM experience impaired health-related quality of life (HRQL), particularly in domains that reflect respiratory symptoms or physical health and activities.[7]

St. George's Respiratory Questionnaire (SGRQ) is a respiratory-specific instrument that was designed to assess HRQL in patients with asthma or COPD.[8] Despite the developer's initial intent—to develop a questionnaire for patients with either of those two conditions—the SGRQ has been used to assess HRQL in patients with other lung diseases, including women with LAM.[9] In fact, investigators have observed that baseline scores from the SGRQ correlate with certain

baseline measures of pulmonary physiology, oxygenation, and functional capacity.[7,9] The SGRQ has been shown to be sensitive to change when used in patients with various respiratory diseases,[10-12] but whether in patients with LAM it can track changes in HRQL that might occur as a result of disease progression or in response to a clinically beneficial (or harmful) medication has never been assessed. An HRQL instrument must possess this attribute to be considered useful as an outcome measure in longitudinal research. We conducted this study to examine the ability of the SGRQ to assess HRQL over time in patients with LAM (ie, to determine its longitudinal construct validity in this disease).

Materials and Methods

We used data collected at baseline, 6 months, and 12 months in the Multicenter International Lymphangioleiomyomatosis Efficacy and Safety of Sirolimus (MILES) trial.[13] The MILES trial was a two-stage trial—a 12-month randomized, double-blinded comparison of sirolimus vs placebo followed by a 12-month observation period—involving 89 patients with LAM who had a $FEV_1 < 70\%$ predicted. The primary end point was the difference between the groups in the rate of change in FEV_1. There were a number of secondary end points, including HRQL as assessed by the SGRQ. We assessed the association between SGRQ scores and certain variables hypothesized to be clinically meaningful measures of LAM severity; henceforth, we refer to those variables as anchors. We hypothesized that changes in the anchors (ie, disease status or severity) would be associated with changes in HRQL and, thus, changes in SGRQ scores.

Saint George's Respiratory Questionnaire

The SGRQ is a self-administered, respiratory-specific questionnaire with three domains (Symptoms, Activity, and Impacts) and a total score designed to assess HRQL.[8] The Symptoms domain, as its name implies, focuses on respiratory symptoms, including breathlessness, cough, and wheeze. The Activity domain probes for physical activities that either cause or are limited by dyspnea. The Impacts domain covers the effects of respiratory disease on several factors, including employment, social interactions, emotional well-being, and the sense of being in control. Scoring weights for the response options for each of the 50 items were derived using data from patients with asthma or COPD.[11] Each domain score and the total score range from 0 to 100, and higher scores connote greater impairment.

Anchors

The four anchors we selected were: (1) FEV_1, (2) diffusing capacity of the lung for carbon monoxide (DLCO), (3) distance walked during the 6-min walk test (6MWD), and (4) serum vascular endothelial growth factor D level (VEGF-D). We chose FEV_1 and DLCO because each has been shown to be impaired in patients with LAM, and as such, they—particularly FEV_1—are measures used universally to characterize LAM severity.[7] Both the FEV_1 and DLCO have been shown in prior cross-sectional studies to correlate with SGRQ scores.[7,9] The 6-min walk test, and 6MWD in particular, is commonly used as a functional assessment in patients with respiratory diseases.[15] We hypothesized that changes in 6MWD would reflect changes in overall physical functionality and that changes in physical functionality would lead to changes in subjects' perceptions of their HRQL. Serum VEGF-D level distinguishes LAM from other diseases, and here, changes in VEGF-D levels were hypothesized to track changes in LAM severity (eg, increases in VEGF-D would be associated with increased LAM severity) and, by association, HRQL.[16]

Statistical Analysis

We performed analyses to examine the association between SGRQ scores and anchor values cross-sectionally, as well as longitudinally, using the data collected at three different time points: baseline, 6 months, and 12 months. To enhance interpretability of results, serum VEGF-D values were log-transformed. In the first set of analyses, Spearman rank correlation was used to assess the relationship among variables cross-sectionally and longitudinally. Linear mixed-effects models were used to further examine these cross-sectional and longitudinal associations simultaneously. For each anchor, we generated four separate mixed-effects models (one model for each of the three SGRQ domains and one for the SGRQ total score). In each model, SGRQ score was the response variable and the anchor was a covariate. Each model included age at enrollment as a time-constant variable, the baseline anchor value (to allow examination of the cross-sectional association between SGRQ score and anchor), and the anchor change from baseline (to allow examination of the longitudinal association). For each model, to account for the within-subject correlation among the repeated measures in SGRQ score, anchor intercept and slope were incorporated as random effects with unstructured covariance. Next, we used general linear models to examine the association between SGRQ changes and quartiles of anchor change (defined as percent change from baseline at 6 or 12 months) after adjusting for the corresponding baseline SGRQ score. Finally, as a visual representation of the relationship between SGRQ change

Manuscript received January 19, 2012; revision accepted November 17, 2012.
Affiliations: From the Autoimmune Lung Center and Interstitial Lung Disease Program (Dr Swigris), National Jewish Health, Denver, CO; the Pediatrics Epidemiology Center (Dr Lee), University of South Florida College of Medicine, Tampa, FL; the Women's Health Research Institute (Dr Cohen), Women's College Hospital, Toronto, ON, Canada; the National Hospital Organization Kinki-Chou Chest Medical Center (Dr Inoue), Osaka, Japan; the National Institutes of Health (Dr Moss), Bethesda, MD; the Lung Transplantation Program (Dr Singer), University of Toronto; the Division of Allergy, Pulmonary, and Critical Care Medicine (Dr Young), Vanderbilt University School of Medicine, Nashville, TN; and the Department of Internal Medicine (Dr McCormack), Division of Pulmonary, Critical Care and Sleep Medicine, University of Cincinnati, Cincinnati, OH.
Funding/Support: Dr Swigris is supported in part by a Career Development Award from the National Institutes of Health (NIH) [Grant K23 HL092227]. Dr Moss was supported by the Intramural Research Program, NIH, National Heart, Lung and Blood Institute. The Multicenter International Lymphangioleiomyomatosis Efficacy and Safety of Sirolimus Trial was supported by the NIH Office of Rare Disease Research [Grant RR019498]; the US Food and Drug Administration [Grant FD003362]; the Japanese Ministry of Health, Labour, and Welfare; the Canadian Institutes of Health Research; Pfizer Pharmaceuticals; The LAM Foundation; the Tuberous Sclerosis Alliance; and Cincinnati Children's Hospital.
Correspondence to: Jeffrey J. Swigris, DO, MS, Associate Professor of Medicine, Autoimmune Lung Center and Interstitial Lung Disease Program, National Jewish Health, Southside Bldg, Office #G011, 1400 Jackson St, Denver, CO 80206; e-mail: swigrisj@njc.org

DOI: 10.1378/chest.12-0161

scores and anchor change scores (again, defined as percent change from baseline), we generated cumulative distribution function (CDF) plots for each SGRQ score using data from subjects in the two extreme quartiles of anchor changes (greatest decline vs greatest improvement). Institutional review board approval was not needed for this study of deidentified, previously collected data. All statistical analyses were performed using SAS, Version 9.2 (SAS Institute Inc.), and $P < .05$ was considered to represent statistical significance for each analysis. Because our analyses were hypothesis driven, we did not adjust for multiple comparisons.

Results

Baseline characteristics of the subjects are displayed in Table 1. On average, airflow limitation was moderately severe. At baseline, 6 months, and 12 months, there were significant correlations—most moderately strong—between various anchor values and SGRQ scores (Table 2). For each anchor, at each time point, the strongest correlations were with Activity domain scores (Table 2). Simple correlations between SGRQ change scores and anchor change scores are presented in Table 3.

The results from the mixed-effects models extend the results from the correlation analyses by yielding estimates for the cross-sectional relationship between SGRQ scores and anchors at baseline (Table 4, "Cross-sectional") as well as how SGRQ scores were predicted to change in relation to changes over time in the anchors (Table 4, "Longitudinal"). At any time point (ie, 6 or 12 months), improvements from baseline in any of the four anchors were predicted to generate improved (lower) SGRQ scores. For example, at 6 or 12 months, every 1% increase in FEV_1% was predicted to yield about a 0.5-point decrease (improved HRQL) in any SGRQ score (domain or total), and a 100-m increase in 6MWD was predicted to yield a 4-point decrease (improved HRQL) in the Symptoms, Activity, or SGRQ total score and a 5-point decrease in Impacts score.

Table 1—*Characteristics of Study Sample*

Variable	Baseline (n = 89)	6 mo (n = 83)	12 mo (n = 74)
Age, y	45.4 ± 10.6	...	...
Race, No. (%)			
White	59 (66)	...	...
Asian	27 (30)	...	...
Other	3 (3)	...	...
Oxygen therapy, No. (%)			
Intermittent use	52 (58)	...	...
Continuous use	28 (31)	...	...
FEV_1			
L	1.37 ± 0.42	1.35 ± 0.42	1.33 ± 0.40
% Predicted	48.5 ± 13.8	48.5 ± 15.0	48.1 ± 14.1
DLCO			
mL/min/mm Hg	10.23 ± 4.61	9.95 ± 4.00	9.62 ± 3.95
% Predicted	43.4 ± 19.0	42.3 ± 16.3	41.2 ± 16.8
6MWD, m	403 ± 105	415 ± 119	425 ± 104
log(VEGF-D), pg/mL	7.22 ± 0.85	7.00 ± 0.82	6.90 ± 0.84
SGRQ scores			
Symptom domain	52.1 ± 18.4	49.8 ± 18.3	48.5 ± 19.2
Activity domain	65.5 ± 17.7	65.2 ± 18.3	64.5 ± 17.6
Impact domain	33.4 ± 17.4	34.5 ± 17.9	31.3 ± 15.0
Total	46.4 ± 15.2	46.6 ± 15.6	44.6 ± 13.7

Data are presented as mean ± SD unless otherwise noted. 6MWD = distance walked during 6-min walk test; DLCO = diffusing capacity of the lung for carbon monoxide; SGRQ = St. George's Respiratory Questionnaire; VEGF-D = serum vascular endothelial growth factor D level.

Table 5 displays mean SGRQ change scores for subgroups stratified on quartiles of change in each anchor. After adjusting for baseline SGRQ score, there were significant differences between various SGRQ scores across quartiles of change in FEV_1, 6MWD, and log(VEGF-D). Although not all relationships were statistically significant, on balance, we observed that greater impairments (ie, decline in FEV_1, decline in DLCO, decline in 6MWD, or increase in VEGF-D) in the anchors were associated with greater impairments (increase) in SGRQ scores, and greater improvement in the anchors were associated with greater improvements (decrease) in SGRQ scores.

The CDF plots (Fig 1) provide a graphical representation of SGRQ change scores for subgroups in the extreme quartiles of change for each anchor. Consider an arbitrarily chosen improvement of at least 10 points in the Symptoms domain (ie, −10 or more extreme on the x-axis in Fig 1A): only 33% of subjects in the first quartile of change in FEV_1 (Q1: subjects with decline in raw FEV_1 of > 11.1% from baseline) compared with 61% of subjects in the fourth quartile of change in FEV_1 (Q4: subjects with an improvement in FEV_1 of > 4.4%) had improvement of at least 10 points in the Symptoms domain from baseline to 12 months. For the Activity domain, 17% in Q1 vs 42% in Q4 had improvement of at least 10 points; for the Impacts domain, 17% in Q1 vs 32% in Q4 had improvement of at least 10 points; and for the SGRQ total, 17% in Q1 vs 28% in Q4 had improvement of at least 10 points.

Discussion

Using data from the MILES trial, we performed several analyses whose results support the SGRQ as an instrument capable of assessing and responding to change over time in HRQL in patients with LAM. When investigators aim to generate data to support the validity of something (for an intended purpose), ideally, there is a gold standard against which it can be measured. Unfortunately, in the realm of HRQL, there is no gold standard. Thus, investigators must choose other variables, hypothesized—or known—to be related to HRQL to serve as anchors. In patients with LAM, pulmonary function tests (FEV_1 in particular)

Table 2—*Spearman Correlation Coefficients Between SGRQ Scores and Anchors at Baseline, 6 Months, and 12 Months*

Domain	Symptom	Activity	Impact	Total
FEV_1%				
Baseline	−0.33 (.002)	−0.35 (.0007)	−0.26 (.01)	−0.33 (.001)
6 mo	−0.41 (.0001)	−0.48 (< .0001)	−0.42 (< .0001)	−0.49 (< .0001)
12 mo	−0.35 (.002)	−0.46 (< .0001)	−0.29 (.01)	−0.40 (.0005)
DLCO%				
Baseline	−0.21 (.048)	−0.43 (< .0001)	−0.20 (.06)	−0.33 (.002)
6 mo	−0.30 (.006)	−0.49 (< .0001)	−0.34 (.002)	−0.45 (< .0001)
12 mo	−0.17 (.14)	−0.54 (< .0001)	−0.33 (.005)	−0.41 (.0004)
6MWD				
Baseline	−0.14 (.19)	−0.36 (.0005)	−0.13 (.23)	−0.23 (.03)
6 mo	−0.28 (.01)	−0.45 (< .0001)	−0.37 (.0006)	−0.46 (< .0001)
12 mo	−0.20 (.08)	−0.49 (< .0001)	−0.37 (.001)	−0.41 (.0003)
Log(VEGF-D)				
Baseline	0.09 (.42)	0.23 (.03)	0.18 (.10)	0.23 (.03)
6 mo	0.21 (.07)	0.27 (.02)	0.10 (.37)	0.19 (.09)
12 mo	0.19 (.12)	0.23 (.06)	0.15 (.20)	0.21 (.08)

Values are Spearman correlation coefficients (*P* value). See Table 1 legend for expansion of abbreviations.

yield important prognostic information and are commonly used to assess disease status. And by extension, they yield some information about patients' well-being or HRQL. Thus, we hypothesized FEV_1 and DLCO as well as 6MWD (as a measure of functional status) would be related to HRQL in patients with LAM and, as such, in the absence of a gold standard for HRQL, would be useful as external anchors for our analyses. Serum VEGF-D has emerged as a diagnostic biomarker for LAM[16]; like FEV_1, DLCO, and 6MWD, we hypothesized VEGF-D levels would change in response to changes in clinically defined disease severity and might, therefore, be associated with HRQL.

The MILES trial revealed that targeting the mTOR (mammalian Target of Rapamycin) pathway is an effective approach to treating LAM, and MILES very likely paved the way for future trials of other mTOR signaling antagonists for LAM. The hope for such a trial is that any physiologic benefit of a drug would translate into improvements in survival and—arguably, perhaps even more importantly—patient symptoms, functional capacity, and sense of well-being. Survival is easy to assess; symptoms and the more abstract constructs (eg, HRQL) are more challenging, but they can be measured, too. The key to doing so is using reliable, valid instruments that are sensitive to underlying change in the construct of interest. Results from the current study support the SGRQ as such an instrument. We identified a number of significant correlations between SGRQ scores and anchors at each of the study time points. The moderately strong correlations we observed (common to analyses like this), rather than being disappointing, are in fact reassuring. They show that the SGRQ performed as hypothesized, but, most importantly, they reveal that the SGRQ captures information about patients with LAM that FEV_1, DLCO, 6MWD, and serum VEGF-D levels do not.

Even more important for the process of building longitudinal validity than identifying significant

Table 3—*Spearman Correlation Coefficients Between SGRQ Change Scores and Anchor Change Scores From Baseline to 6 and 12 Months*

Domain	Symptom	Activity	Impact	Total
FEV_1%				
6 mo	−0.18 (.09)	−0.34 (.0016)	−0.37 (.0005)	−0.42 (< .0001)
12 mo	−0.16 (.18)	−0.16 (.1621)	−0.24 (.0378)	−0.27 (.0204)
DLCO%				
6 mo	−0.02 (.83)	0.07 (.51)	−0.08 (.46)	−0.05 (.65)
12 mo	−0.04 (.76)	−0.11 (.32)	−0.05 (.66)	−0.08 (.52)
6MWD				
6 mo	−0.07 (.51)	−0.17 (.1181)	−0.15 (.18)	−0.18 (.1128)
12 mo	−0.22 (.06)	−0.30 (.0086)	−0.22 (.06)	−0.35 (.0023)
Log(VEGF-D)				
6 mo	0.06 (.63)	0.09 (.45)	0.19 (.09)	0.20 (.09)
12 mo	0.10 (.39)	0.15 (.23)	0.09 (.46)	0.15 (.23)

Values are Spearman correlation coefficients (*P* value). See Table 1 legend for expansion of abbreviations.

Table 4—*Mixed-Effects Model-Generated Parameter Estimates for SGRQ Change Scores Resulting From Change in Anchor Scores*

Associations	Domain	Symptoms	Activity	Impact	Total
Cross-sectional	FEV_1, mL	−0.014 ± 0.004 (.002)	−0.016 ± 0.004 (< .0001)	−0.015 ± 0.004 (.0009)	−0.015 ± 0.004 (.0001)
	FEV_1%	−0.46 ± 0.13 (.0009)	−0.48 ± 0.11 (< .0001)	−0.39 ± 0.13 (.003)	−0.42 ± 0.11 (.0002)
	DLCO, mL/min/mm Hg	−1.11 ± 0.45 (.02)	−1.72 ± 0.40 (< .0001)	−0.84 ± 0.58 (.21)	−1.15 ± 0.38 (.005)
	DLCO %	−0.30 ± 0.11 (.01)	−0.42 ± 0.10 (< .0001)	−0.20 ± 0.11 (.08)	−0.29 ± 0.09 (.004)
	6MWD, m	−0.03 ± 0.02 (.12)	−0.07 ± 0.02 (.0002)	−0.04 ± 0.02 (.03)	−0.05 ± 0.02 (.004)
	Log(VEGF-D)	4.14 ± 2.16 (.06)	5.95 ± 2.22 (.009)	3.45 ± 2.17 (.12)	4.34 ± 1.95 (.03)
Longitudinal	FEV_1, mL	−0.018 ± 0.007 (.02)	−0.020 ± 0.005 (< .0001)	−0.021 ± 0.006 (.0007)	−0.020 ± 0.005 (< .0001)
		0.583	0.744	0.714	0.750
	FEV_1%	−0.49 ± 0.22 (.03)	−0.56 ± 0.14 (< .0001)	−0.61 ± 0.18 (.001)	−0.57 ± 0.14 (.0003)
		0.531	0.759	0.744	0.776
	DLCO, mL/min/mm Hg	−0.23 ± 0.75 (.76)	−0.42 ± 0.56 (.47)	−0.03 ± 0.82 (.97)	−0.28 ± 0.53 (.60)
		0.580	0.788	0.784	0.797
	DLCO %	−0.06 ± 0.18 (.73)	−0.10 ± 0.14 (.49)	−0.05 ± 0.15 (.76)	−0.09 ± 0.12 (.47)
		0.580	0.798	0.783	0.803
	6MWD, m	−0.05 ± 0.02 (.01)	−0.02 ± 0.02 (.15)	−0.05 ± 0.02 (.008)	−0.04 ± 0.01 (.003)
		0.612	0.789	0.716	0.776
	Log(VEGF-D)	4.85 ± 2.51 (.06)	1.63 ± 1.70 (.34)	3.86 ± 1.89 (.04)	3.27 ± 1.51 (.03)
		0.636	0.825	0.763	0.815

Values are mixed-effects model parameter estimate ± SE (*P* value). In the Longitudinal section of the table, the value below the *P* value is the lowest of the estimated within-patient correlations from the model. See Table 1 for expansion of abbreviations.

correlations between static anchor values and SGRQ scores, we found that SGRQ scores tracked changes in each of the four anchors over time. For example, from the mixed-effects analyses, a 200-mL increase in FEV_1, or an 8% increase in FEV_1%, was predicted to result in a 4-point decline in SGRQ total score (connoting an improvement in HRQL). Likewise, a 5% increase in DLCO% was predicted to result in a >5-point decline in SGRQ Activity score. These analyses also revealed that change in VEGF-D was significantly associated with change in SGRQ scores, but this was for very large (one log) changes in serum VEGF-D levels. Given what is known about serum VEGF-D, this was not too surprising to us: VEGF-D is a diagnostic marker, but its value as a biomarker of pulmonary disease activity (i.e., whether longitudinal changes in VEGF-D correlate with lung disease progression in LAM) is, at present, uncertain. Perhaps

Table 5—*SGRQ Change Scores for Subjects Stratified Into Quartiles Based on Change From Baseline to 12 Months in Anchor Values*

Anchor	Symptoms	*P* Value	Activity	*P* Value	Impacts	*P* Value	Total	*P* Value
FEV_1		**.01**		**.006**		**.16**		**.01**
≤ −11.2% (Q1)	−0.12 ± 14.6	.017	−1.71 ± 8.76	.06	0.95 ± 12.1	.025	−0.13 ± 7.98	.006
−11.2% to −3.4% (Q2)	1.45 ± 13.1	.015	1.31 ± 9.84	.004	−0.34 ± 9.96	.17	0.34 ± 8.16	.01
−3.4% to 4.4% (Q3)	4.57 ± 21.4	.002	3.26 ± 8.43	.001	−2.30 ± 13.1	.24	0.66 ± 10.0	.007
>4.4% (Q4)	−13.4 ± 14.6	Ref	−6.46 ± 8.20	Ref	−6.25 ± 12.8	Ref	−7.53 ± 10.0	Ref
DLCO		**.6**		**.5**		**.6**		**.6**
≤ −9.9% (Q1)	−3.74 ± 15.2	.736	0.65 ± 8.57	.418	−1.83 ± 13.1	.363	−1.29 ± 8.42	.388
−9.9% to −4.7% (Q2)	2.27 ± 18.5	.193	0.06 ± 8.28	.559	−2.06 ± 10.1	.582	−0.82 ± 8.58	.402
−4.7% to 4.7% (Q3)	−1.02 ± 15.9	.438	−3.14 ± 9.75	.522	−3.15 ± 11.7	.687	−2.92 ± 9.98	.789
>4.7% (Q4)	−4.60 ± 20.1	Ref	−1.04 ± 11.1	Ref	−2.20 ± 13.9	Ref	−2.07 ± 11.8	Ref
6MWD		**.7**		**.02**		**.17**		**.04**
≤ −2.3% (Q1)	−1.03 ± 17.9	.400	2.87 ± 7.73	.05	1.81 ± 11.9	.233	2.02 ± 9.18	.06
−2.3% to 4.5% (Q2)	2.54 ± 22.7	.241	0.97 ± 6.53	.37	−0.75 ± 9.35	.662	0.34 ± 7.99	.36
4.5% to 17.1% (Q3)	−3.32 ± 14.5	.449	−5.41 ± 10.8	.26	−6.56 ± 14.8	.301	−5.88 ± 10.2	.33
>17.1% (Q4)	−5.53 ± 11.9	Ref	−2.22 ± 10.7	Ref	−2.60 ± 11.4	Ref	−2.88 ± 9.41	Ref
Log(VEGF-D)		**.04**		**.8**		**.5**		**.16**
≤ −9.0% (Q1)	−7.54 ± 15.3	.7	−2.22 ± 8.43	.602	−5.97 ± 16.1	.452	−5.17 ± 11.6	.437
−9.0% to −3.7% (Q2)	−2.70 ± 18.4	.4	−2.48 ± 9.75	.859	−0.78 ± 11.6	.628	−1.51 ± 10.1	.440
−3.7% to 0.8% (Q3)	6.29 ± 17.9	.02	1.43 ± 9.65	.625	2.76 ± 11.5	.505	2.95 ± 9.03	.158
>0.8% (Q4)	−4.88 ± 12.8	Ref	0.14 ± 10.8	Ref	−3.27 ± 8.49	Ref	−2.44 ± 5.86	Ref

Values are mean ± SD. Boldface numbers are *P* values for comparison across SGRQ means within anchor; other *P* values correspond to pairwise comparisons of SGRQ means with anchor using Q4 as the reference. See Table 1 legend for expansion of abbreviations.

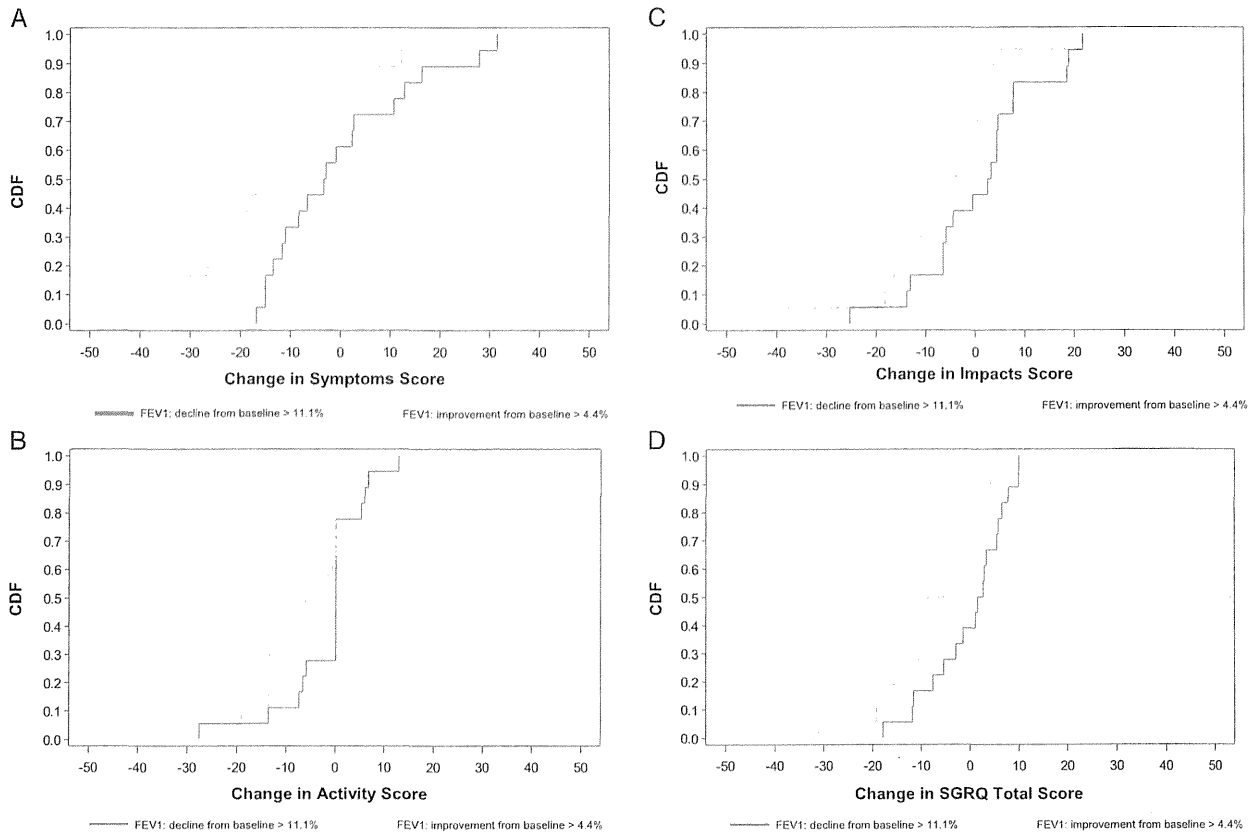

FIGURE 1. A, Plot of CDF for the two extreme quartiles of change from baseline in FEV_1 for SGRQ Symptoms domain. B, Plot of CDF for the two extreme quartiles of change from baseline in FEV_1 for SGRQ Activity domain. C, Plot of CDF for the two extreme quartiles of change from baseline in FEV_1 for SGRQ Impacts domain. D, Plot of CDF for the two extreme quartiles of change from baseline in FEV_1 for SGRQ total score. CDF = cumulative distribution function; SGRQ = St. George's Respiratory Questionnaire.

VEGF-D levels are exquisitely sensitive to LAM activity (at least lymphatic involvement) at the molecular level, but our clinical metrics (eg, FEV_1, DLCO, and 6MWD) are not: only extensive and prolonged lymphatic involvement (that drives up VEGF-D) translates into clinical worsening able to be captured by FEV_1, DLCO, or 6MWD. Clearly, more research into the role of VEGF-D as a biomarker of disease activity in patients with established LAM is warranted.

An effect of being sensitive to longitudinal changes in HRQL within a population is the ability to distinguish change over time between subgroups in that population. When we grouped subjects according to change over time in FEV_1, on balance, the SGRQ performed as hypothesized: subjects with the greatest improvement in FEV_1 experienced the greatest improvement in HRQL. For that particular set of analyses, as has been noted previously in similar analyses for another respiratory disease,[12] DLCO presents challenges as an anchor. First, DLCO changed very little (average change < 1 mL/mm Hg/min among all subjects) over the duration of MILES. Second, DLCO is a statistically noisy variable, and changes in DLCO may fluctuate for reasons unrelated to changes in LAM severity. For example, DLCO is affected by changes in hemoglobin; DLCO was not adjusted for hemoglobin in MILES. Also, DLCO results are often affected by maldistribution of the inspired test gas mixture. This can occur when residual volume is elevated; in MILES, average residual volume was 141% of the predicted value. Finally, for DLCO and the other anchors, the loss of power induced by categorizing variables (as we did in this set of analyses) likely contributed to the inability to identify statistically significant differences between subgroups.

The SGRQ is just one of many instruments designed to assess HRQL and one of several respiratory-specific tools developed for this purpose. A mistake that has been perpetuated in the medical literature is that one cross-sectional correlation study—if statistically significant results emerge—"validates" an instrument for use in a longitudinal trial.[17] In a handful of studies, investigators have generated data to support this so-called "concurrent validity" for the SGRQ (and other instruments) in LAM,[7,9] but, to our knowledge, ours is the first to assess whether the SGRQ can be used confidently in LAM to assess change in longitudinal studies (eg, drug trials).[18-20] The results of our

analyses are not surprising; however, these analyses must be done to confirm the SGRQ performs as hypothesized in LAM. This is the essence of building validity. The FDA has formalized recommendations for how instruments like the SGRQ might qualify as a valid, reliable outcome measure whose scores have "interpretable meaning" in a target population.[21] To our knowledge, there have been no HRQL questionnaires submitted to the FDA for LAM, but data from this study could be useful in such a submission.

This study has limitations: the first is that subjects in the MILES trial may not be representative of the general LAM population. Given how uncommon LAM is, we doubt this is the case. However, subjects in MILES had moderately severe airflow limitation—more severe than the cohort enrolled in a nationwide registry[7]—so whether the SGRQ would perform equally well in a trial that included only subjects with milder LAM is uncertain. The results of our study cannot be extended to other HRQL questionnaires. It is possible that other HRQL questionnaires would perform similarly (or even better than the SGRQ) in LAM under the same circumstances. Until their longitudinal construct validity is assessed, those instruments cannot be used confidently in longitudinal LAM research. On the face of it, the SGRQ contains many items—those that ask about wheeze, cough, dyspnea, and physical activities—relevant to LAM patients.[7] However, the SGRQ was not developed specifically for patients with LAM. Whether an instrument developed by specifically incorporating perspectives from patients with LAM would perform better than the SGRQ is unknown. Although not necessarily relevant to this study, in certain studies of patients with COPD, women's scores are higher (greater impairment in HRQL) than men's for the SGRQ.[22] Some readers may not be familiar with mixed-effects models, but they are believed by most experts to be the models of choice when analyzing longitudinal data (including HRQL data from therapeutic trials).[23] These models parameterize the within-subject correlation that results from repeatedly measuring an outcome over time in the same individual, and they accommodate both incomplete data and time-varying covariates. In contrast, when simple correlation (or even linear regression) is used to assess the relationship between two variables, if data for either variable are missing for a subject, by necessity the subject is deleted from the analysis. Thus, mixed-effects models provide the most statistically efficient method for analyzing longitudinal data. Finally, certain analyses did not yield statistically significant results, but this was largely due to a loss of power resulting from categorization of continuous variables (eg, those in which the anchors were stratified into quartiles). The majority of our analyses support the longitudinal validity of the SGRQ in LAM.

Acknowledgments

Author contributions: Dr Swigris had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Dr Swigris: contributed to study conceptualization and planning, generating intellectual content for the manuscript, and critiquing and approving final content.

Dr Lee: contributed to study conceptualization and planning, analyzing data, generating intellectual content for the manuscript, and critiquing and approving final content.

Dr Cohen: contributed to generating intellectual content for the manuscript and critiquing and approving final content.

Dr Inoue: generating intellectual content for the manuscript and critiquing and approving final content.

Dr Moss: generating intellectual content for the manuscript and critiquing and approving final content.

Dr Singer: generating intellectual content for the manuscript and critiquing and approving final content.

Dr Young: generating intellectual content for the manuscript and critiquing and approving final content.

Dr McCormack: contributed to study conceptualization and planning, generating intellectual content for the manuscript, and critiquing and approving final content.

Financial/nonfinancial disclosures: The authors have reported to *CHEST* that no potential conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article.

Role of sponsors: The sponsors had no role in the design of the study, the collection and analysis of the data, or in the preparation of the manuscript.

Other contributions: We thank Margaret Bevans, RN, PhD, Clinical Center, National Institutes of Health, for critical reading of the manuscript.

References

1. Cabana H, Alexandre C, Agathos SN, Jones JP. Immobilization of laccase from the white rot fungus Coriolopsis polyzona and use of the immobilized biocatalyst for the continuous elimination of endocrine disrupting chemicals. *Bioresour Technol.* 2009;100(14):3447-3458.
2. Corrin B, Liebow AA, Friedman PJ. Pulmonary lymphangiomyomatosis. A review. *Am J Pathol.* 1975;79(2):348-382.
3. Taylor JR, Ryu J, Colby TV, Raffin TA. Lymphangioleiomyomatosis. Clinical course in 32 patients. *N Engl J Med.* 1990; 323(18):1254-1260.
4. Johnson SR, Tattersfield AE. Decline in lung function in lymphangioleiomyomatosis: relation to menopause and progesterone treatment. *Am J Respir Crit Care Med.* 1999;160(2):628-633.
5. Taveira-DaSilva AM, Stylianou MP, Hedin CJ, Hathaway O, Moss J. Decline in lung function in patients with lymphangioleiomyomatosis treated with or without progesterone. *Chest.* 2004;126(6):1867-1874.
6. Johnson SR, Whale CI, Hubbard RB, Lewis SA, Tattersfield AE. Survival and disease progression in UK patients with lymphangioleiomyomatosis. *Thorax.* 2004;59(9):800-803.
7. Ryu JH, Moss J, Beck GJ, et al; NHLBI LAM Registry Group. The NHLBI lymphangioleiomyomatosis registry: characteristics of 230 patients at enrollment. *Am J Respir Crit Care Med.* 2006;173(1):105-111.
8. Jones PW, Quirk FH, Baveystock CM. The St George's Respiratory Questionnaire. *Respir Med.* 1991;85(suppl B):25-31.
9. Xu KF, Wang L, Tian XL, et al. The St. George's Respiratory Questionnaire in lymphangioleiomyomatosis. *Chin Med Sci J.* 2010;25(3):140-145.

10. Jones PW. Interpreting thresholds for a clinically significant change in health status in asthma and COPD. *Eur Respir J*. 2002;19(3):398-404.
11. Spencer S, Calverley PM, Sherwood Burge P, Jones PW; ISOLDE Study Group. Inhaled Steroids in Obstructive Lung Disease. Health status deterioration in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2001;163(1):122-128.
12. Swigris JJ, Brown KK, Behr J, et al. The SF-36 and SGRQ: validity and first look at minimum important differences in IPF. *Respir Med*. 2010;104(2):296-304.
13. McCormack FX, Inoue Y, Moss J, et al; National Institutes of Health Rare Lung Diseases Consortium; MILES Trial Group. Efficacy and safety of sirolimus in lymphangioleiomyomatosis. *N Engl J Med*. 2011;364(17):1595-1606.
14. Jones PW, Quirk FH, Baveystock CM, Littlejohns P. A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. *Am Rev Respir Dis*. 1992;145(6):1321-1327.
15. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med*. 2002; 166(1):111-117.
16. Young LR, Vandyke R, Gulleman PM, et al. Serum vascular endothelial growth factor-D prospectively distinguishes lymphangioleiomyomatosis from other diseases. *Chest*. 2010; 138(3):674-681.
17. Revicki DA, Osoba D, Fairclough D, et al. Recommendations on health-related quality of life research to support labeling and promotional claims in the United States. *Qual Life Res*. 2000;9(8):887-900.
18. Streiner D, Norman G. *Health Measurement Scales: A Practical Guide to Their Development and Use*. 4th ed. New York, NY: Oxford University Press; 2008.
19. Hays RD, Hadorn D. Responsiveness to change: an aspect of validity, not a separate dimension. *Qual Life Res*. 1992; 1(1):73-75.
20. Patrick DL, Chiang YP. Measurement of health outcomes in treatment effectiveness evaluations: conceptual and methodological challenges. *Med Care*. 2000;38(suppl 9):II14-II25.
21. Food and Drug Administration Center for Drug Evaluation and Research. *Guidance for Industry: qualification process for drug development tools*. Silver Spring, MD: Office of Communications, Division of Drug Information, Center for Drug Evaluation and Research, Food and Drug Administration; 2010.
22. de Torres JP, Casanova C, Hernández C, et al. Gender associated differences in determinants of quality of life in patients with COPD: a case series study. *Health Qual Life Outcomes*. 2006;4:72.
23. Fairclough D. *Design and Analysis of Quality of Life Studies in Clinical Trials*. 2nd ed. New York, NY: CRC Press; 2010.

Int J Clin Oncol (2013) 18:743–749
DOI 10.1007/s10147-012-0438-0

ORIGINAL ARTICLE

Drug-induced lung injury associated with sorafenib: analysis of all-patient post-marketing surveillance in Japan

Yuka Horiuchi-Yamamoto · Akihiko Gemma · Hiroyuki Taniguchi · Yoshikazu Inoue · Fumikazu Sakai · Takeshi Johkoh · Kiminori Fujimoto · Shoji Kudoh

Received: 14 March 2012 / Accepted: 7 June 2012 / Published online: 30 June 2012

Abstract
Background Sorafenib is a multi-kinase inhibitor currently approved in Japan for unresectable and/or metastatic renal cell carcinoma and unresectable hepatocellular carcinoma. Although drug-induced lung injury has recently been the focus of interest in Japanese patients treated with molecular targeting agents, the clinical features of patients receiving sorafenib remain to be completely investigated.
Methods All-patient post-marketing surveillance data was obtained within the frame of Special Drug Use Investigation; between April 2008 and March 2011, we summarized the clinical information of 62 cases with drug-induced lung injury among approximately 13,600 sorafenib-treated patients in Japan. In addition, we summarized the results of evaluation by a safety board of Japanese experts in 34 patients in whom pulmonary images were available. For the calculation of reporting frequency, interim results of Special Drug Use Investigation were used.
Results In the sets of completed reports (2,407 in renal cell carcinoma and 647 in hepatocellular carcinoma), the reporting frequency was 0.33 % (8 patients; fatal, 4/8) and 0.62 % (4 patients; fatal, 2/4), respectively. Major clinical symptoms included dyspnea, cough, and fever. Evaluation of the images showed that 18 cases out of 34 patients had a pattern of diffuse alveolar damage. The patients with hepatocellular carcinoma showed a greater incidence and earlier onset of lung injury than those with renal cell carcinoma.
Conclusion Although the overall reporting frequency of sorafenib-induced lung injury is not considered high, the radiological diffuse alveolar damage pattern led to a fatal

Y. Horiuchi-Yamamoto (✉)
Pharmacovigilance, Medical Affairs, Bayer Yakuhin, Ltd.,
4-9 Umeda 2-chome, Kita-ku, Osaka 530-0001, Japan
e-mail: yuka.yamamoto@bayer.com

A. Gemma
Department of Pulmonary Medicine/Infection and Oncology,
Internal Medicine, Nippon Medical School,
Tokyo 113-8603, Japan

H. Taniguchi
Department of Respiratory Medicine and Allergy,
Tosei General Hospital, Seto, Aichi 489-8642, Japan

Y. Inoue
Department of Diffuse Lung Diseases and Respiratory Failure,
Clinical Research Center, National Hospital Organization Kinki-Chuo Chest Medical Center, Sakai, Osaka 591-8555, Japan

F. Sakai
Department of Diagnostic Radiology, International Medical Center, Saitama Medical University, Hidaka,
Saitama 350-1298, Japan

T. Johkoh
Department of Radiology, Kinki Central Hospital of Mutual Aid Association of Public School Teachers,
Itami, Hyogo 664-8533, Japan

K. Fujimoto
Department of Radiology, Kurume University
School of Medicine, and Center for Diagnostic Imaging,
Kurume University Hospital, Kurume, Fukuoka 830-0011, Japan

S. Kudoh
Fukujuji Hospital, Japan Anti-tuberculosis Association as Public Interest Incorporated Foundation, Kiyose,
Tokyo 204-8522, Japan

outcome. Therefore, early recognition of sorafenib-induced lung injury is crucial for physicians and patients.

Keywords Sorafenib · Drug-induced lung injury · Interstitial lung disease · Drug-related adverse event · Japanese · Post-marketing surveillance

Introduction

Sorafenib is a small-molecule multikinase inhibitor targeting several serine/threonine and receptor tyrosine kinases and interacts with multiple intracellular (CRAF, BRAF, and mutant BRAF) and cell surface (KIT, FLT-3, VEGFR-2, VEGFR-3, and PDGFR-β) kinases [1–3]. Sorafenib has been approved for renal cell and hepatocellular carcinoma [4, 5], and clinical studies are in progress for several other types of tumors. Sorafenib was approved in Japan for unresectable and/or metastatic renal cell carcinoma (RCC) in 2008 and for unresectable hepatocellular carcinoma (HCC) in 2009. The major drug-related adverse events (AE) of sorafenib include hand-foot skin reaction, diarrhea, hypertension, and increased pancreatic enzyme levels [4, 5]. Drug-induced lung injury (DLI) was added to the Japanese package insert of sorafenib, along with the issuance and distribution of "Safety information for acute lung injury/interstitial pneumonia" in December 2008 [6], and close monitoring of patients has been ongoing since then.

Chemotherapeutic drugs that most commonly cause DLI include paclitaxel, docetaxel, gemcitabine, and irinotecan [7, 8]. DLI in Japanese patients treated with molecular targeting agents has been the focus of many studies [9]. Among tyrosine kinase inhibitors, gefitinib and erlotinib are associated with an increase in the incidence of DLI in Japanese patients [10–12]. The precise incidence and clinical characteristics of DLI associated with sorafenib have not been reported although the cases of some individual patients have been reported [13]. Here, we investigated the clinical features of Japanese patients with sorafenib-associated lung injury in a post-marketing surveillance setting.

Patients and methods

Patients

The patient flow of the surveillance is shown in Fig. 1. Between April 2008 and March 2011, sorafenib was administered to approximately 13,600 patients (approximately 5,500 for RCC and 8,100 for HCC) within the frame of Special Drug Use Investigation (SDUI). SDUI is a post-marketing surveillance method specific to Japan, performed under the instruction of the Japanese Health Authority (Pharmaceuticals and Medical Device Agency [PMDA]) for investigating safety and efficacy of sorafenib in clinical practice. For sorafenib, an SDUI with all-patient investigation system was required. The institutions participating in the SDUI concluded a contract with Bayer Yakuhin, Ltd. Sixty-two patients with DLI were identified during the above period on the basis of either an AE reported by a physician or by an evaluation of image by a board of experts ("Safety Advisory Board for interstitial lung disease in Nexavar®"; hereafter, "the ILD Ad-board"). For the calculation of reporting frequency, the interim SDUI results of 2,407 patients with RCC and 647 patients with HCC were used [14, 15]. Reported terms were encoded to the Medical Dictionary for Regulatory Activities Preferred Terms (MedDRA-PT).

Collection of clinical information

Clinical information was collected for the initial purpose of reporting AE from Bayer Yakuhin, Ltd. to the Japanese Health Authority.

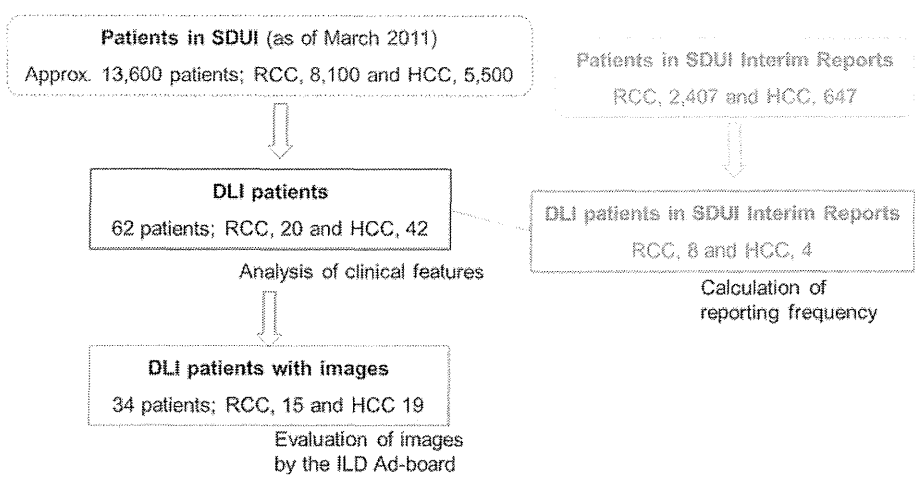

Fig. 1 Breakdown of patients according to SDUI and DLI. *DLI* drug-induced lung injury, *HCC* hepatocellular carcinoma, *ILD Ad-board* ILD Safety Advisory Board, *RCC* renal cell carcinoma, *SDUI* Special Drug Use Investigation

Table 1 Demographics of patients with DLI

	RCC $n = 20$	HCC $n = 42$	Total $n = 62$
Age (years)			
Median (range)	71.5 (50–83)	70 (51–84)	70 (50–84)
Gender			
Male/female	12/8	34/8	46/16
Performance status (ECOG-PS)			
0, 1/2 or over/unknown	16/0/4	37/2/3	53/2/7
Concomitant disease or past history of pulmonary disease			
Chronic type of interstitial pneumonia/pulmonary fibrosis	5	4	9
Other pulmonary disease	2	2	4

ECOG-PS Eastern Cooperative Oncology Group Performance Status

For all patients with DLI, reporting physicians were requested to respond to a special questionnaire. This questionnaire included detailed clinical course, laboratory data and reports of image evaluation, as well as basic information for AE reporting such as patient age, gender, past medical history, concomitant diseases, treatment duration of sorafenib, and concomitant medications. In addition, reporting physicians were requested to provide chest radiograph and/or chest computed tomography (CT) scan. "Fatal" cases included those patients in whom the outcome of DLI was reported as fatal by reporting physicians.

Image interpretation

Imaging data was provided by reporting physicians in the case of 34 patients. In 33 of these patients, conventional CT and/or high-resolution CT were used for evaluation at onset, and in 1 patient, only chest radiograph was used. These images were evaluated as digitized files in Digital Imaging and Communication in Medicine (DICOM) format with anonymization. The members of the ILD Ad-board were requested to independently evaluate these imaging data with DICOM viewer software by referring to clinical information according to a specific evaluation sheet, including preexisting pulmonary conditions, ILD scoring, and image pattern of DLI. ILD scoring indicates the compatibility with ILD in 5 levels, from the score 1 with the finding that ILD can be denied to the score 5 with the finding that ILD is definite. The image patterns were classified as being most consistent with 1 of 2 patterns: diffuse alveolar damage (DAD) or non-DAD [16, 17]. Image patterns demonstrating extensive bilateral ground-glass opacities with or without consolidation and/or architectural distortion, or features suggestive of fibrosis in a predominantly depending distribution were considered as DAD [18, 19]. Case review and final decision of image interpretation were made periodically in a panel discussion including all members of the ILD Ad-board. The conclusions of the ILD Ad-board were conveyed to each reporting physician and to the Japanese Health Authority.

Results

Patient demographics

Patient demographics are listed in Table 1. Among 62 patients with DLI, 20 had RCC and 42 had HCC as underlying disease. Concomitant disease and medical history included chronic type of interstitial pneumonia/pulmonary fibrosis in 9 patients. Other pulmonary diseases included bronchial asthma, emphysema, pleuritis, and history of surgical intervention to the lung.

Reporting frequency

The reporting frequency of DLI was 0.33 % (8/2,407, 4 out of 8 patients had fatal outcome of DLI) for RCC and 0.62 % (4/647, 2 out of 4 patients had fatal outcome of DLI) for HCC according to respective interim reports of SDUI on the basis of clarification of precise number of patients treated with sorafenib. The 8 patients with RCC and 4 patients with HCC are included in the 62 DLI patients among approximately 13,600 sorafenib-treated patients. The reporting frequency described above from interim results of SDUI is considered to be equivalent to the overall accumulation status of DLI patients as of March 2011: 20 among approximately 5,500 patients with RCC, and 42 among approximately 8,100 patients with HCC.

Time to onset

Time to onset, the interval between the start of administration and the onset of DLI, is shown in Fig. 2. Overall,

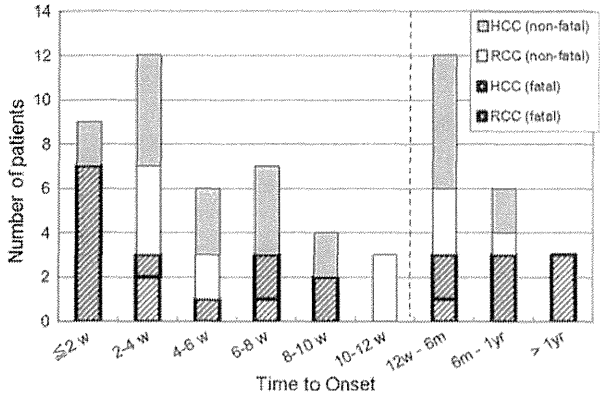

Fig. 2 Time to onset for DLI. *White bars* indicate patients with renal cell carcinoma. *Gray bars* indicate patients with hepatocellular carcinoma. *White shadow* or *gray shadow bars* indicate patients with renal cell carcinoma or hepatocellular carcinoma with a fatal outcome, respectively

Table 2 Treatment for DLI

Treatment for DLI	Number of patients (fatal)		
	RCC $n = 20$	HCC $n = 42$	Total $n = 62$
Steroid pulse therapy	8 (3)	18 (10)	26 (13)
Start on the day or the next day of onset	4 (1)	15 (10)	19 (11)
Start subsequently	4 (2)	1 (0)	5 (2)
No information about the timing of initiation	0 (0)	2 (0)	2 (0)
Other steroid (except pulse therapy)	2 (1)	7 (3)	9 (4)
Other medication (except steroid)	0 (0)	4 (3)	4 (3)
No medication	5 (1)	3 (1)	8 (2)
No information about treatment for DLI	5 (2)	10 (1)	15 (3)
Total	20 (7)	42 (18)	62 (25)

the peak time to onset in all patients was during 2–4 weeks after the start of administration; however, in some patients, DLI occurred more than 6 months after the start of administration. The 9 patients with time to onset within 2 weeks were all HCC patients. Furthermore, a tendency of earlier onset was observed in HCC patients (median 50 days, range 2–289 days) than in RCC patients (median 74 days, range 16–420 days).

Clinical symptoms

Data about clinical signs and symptoms were available in 47 of the 62 patients. Among these 47 patients, dyspnea, cough, and fever, were frequently reported symptoms

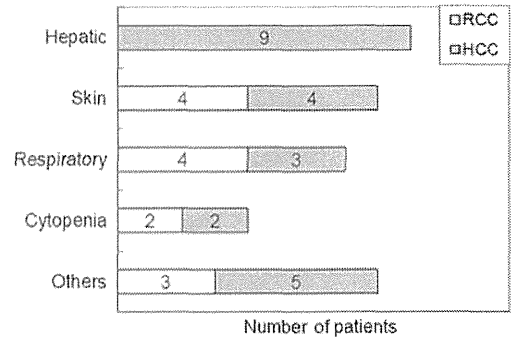

Fig. 3 Other serious drug-related adverse events of sorafenib in patients with DLI

Table 3 Image evaluation by the ILD Ad-board

Image pattern	Number of patients (fatal)		
	RCC $n = 20$	HCC $n = 42$	Total $n = 62$
DAD	8 (5)	10 (7)	18 (12)
Non-DAD	7 (1)	8 (2)	15 (3)
OP	3 (0)	1 (0)	4 (0)
Others	2 (1)	7 (2)	9 (3)
Pre-existing ILD	2 (0)	0 (0)	2 (0)
ILD excluded	0 (0)	1 (0)	1 (0)
Image not available	5 (1)	23 (9)	28 (10)
Total	20 (7)	42 (18)	62 (25)

Pre-existing ILD included asbestosis and chronic type of interstitial pneumonia

DAD diffuse alveolar damage, *ILD* interstitial lung disease, *OP* organizing pneumonia

(in 34, 20, and 15 patients, respectively). Hemoptysis was observed in 2 patients. Five patients were asymptomatic.

Treatment for DLI

Treatment for DLI is shown in Table 2. Among 47 patients for whom information about DLI treatment was provided, 35 received steroid administration, including steroid pulse therapy (intravenous high-dose steroid therapy) in 26 patients. Of the 26 patients receiving steroid pulse therapy, 19 were treated with pulse therapy on the day or the next day of the onset. Medications besides steroids included antibiotics and neutrophil elastase inhibitors.

Other serious drug-related AEs of sorafenib in patients with DLI

Sorafenib-related serious AEs reported in addition to DLI are shown in Fig. 3. Hepatic disorder was the most

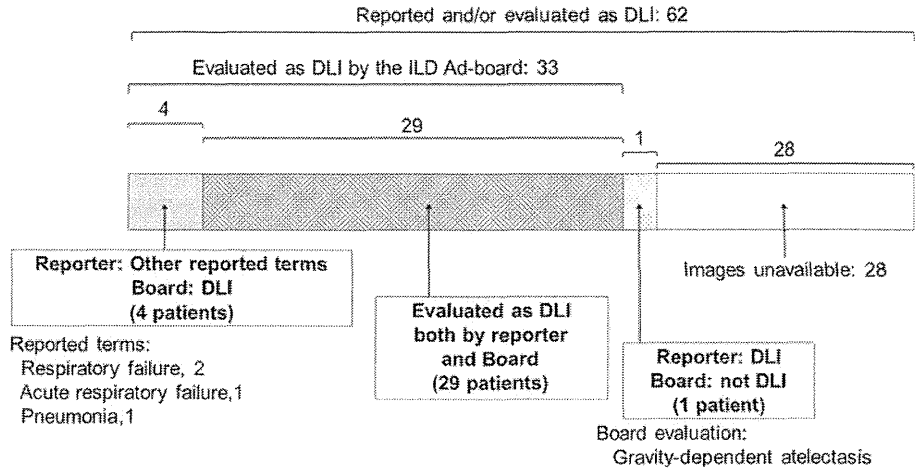

Fig. 4 Relation between the reported terms by reporting physicians and the result of image evaluation by the ILD Ad-board. *DLI* drug-induced lung injury, *ILD Ad-board* ILD Safety Advisory Board

frequently reported AE in 9 patients with HCC. Skin disorder and cytopenic events occurred in 8 and 4 patients, respectively. Respiratory disorders such as dyspnea or respiratory failure, which are the symptoms of DLI itself, were redundantly reported.

Imaging findings

The results of image evaluation are shown in Table 3. The images of 18 patients showed DAD pattern and 15 showed non-DAD pattern. Twelve out of 18 patients with DAD pattern had a fatal outcome. In contrast, 3 out of 15 patients with non-DAD pattern had a fatal outcome.

The relationship between the terms reported by reporting physicians and the result of image evaluation by the ILD Ad-board are shown in Fig. 4. In the cases of 29 patients, the terms reported by the reporting physicians and the result of the evaluation by the ILD Ad-board were consistent. The condition of 4 patients, ultimately diagnosed as DLI by the ILD Ad-board, had been reported initially by reporting physicians as respiratory failure (2 patients), acute respiratory failure (1 patient), and pneumonia (1 patient). The condition of 1 patient diagnosed as DLI by a reporting physician was excluded by the ILD Ad-board and was determined to be gravity-dependent atelectasis.

Discussion

Drug-induced lung injury, especially caused by novel anticancer drugs, has recently been the focus of many studies. Gefitinib-induced DLI is 3.5 % in a retrospective analysis [11] and 5.8 % in a prospective study [10] of Japanese patients with non-small cell lung cancer (NSCLC). In a cohort study, including gefitinib and chemotherapy in Japanese patients with NSCLC, the naive cumulative incidence rates at the end of 12-week follow-up were 4.0 % for gefitinib versus 2.1 % for conventional chemotherapy [20]. Another study in Japanese patients with NSCLC showed that the incidence of DLI within the first month of treatment was 1.0 % for erlotinib versus 2.4 % for gefitinib [12]. Although the reporting frequency of sorafenib-induced lung injury in RCC and HCC patients in the present analysis is not considered high, the difference in underlying malignancy between pulmonary and non-pulmonary origin should be taken into consideration.

Several reports indicate that Japanese patients are more likely to develop DLI. The worldwide prevalence of DLI in gefitinib-treated patients was approximately 1 %, and that in a US manufacturer expanded access program was 0.3 % [21]. One of the reasons for this difference is considered to be solicitation bias due to a difference in the system for collecting information about AE in the post-marketing setting. In Japan, active solicitation is required in the early phase of launch of new drugs under the Early Post-marketing Phase Vigilance (EPPV) requirement of the Japanese Health Authority, whereas worldwide reporting frequency is basically estimated from spontaneous reports or literature reports. In addition, the high incidence of DLI in Japan may be because of the greater awareness about DLI. Genetic susceptibility to DLI is also suggested; however, further research is required to address this issue [22].

The relation between underlying malignancy and DLI was that HCC patients tended to develop DLI earlier than RCC patients, and hepatic disorder was the most frequently reported AE other than DLI in HCC patients. The reason for this remains to be completely elucidated; however, several factors are postulated, such as impaired metabolism of sorafenib in patients with decreased hepatic function reserve and any patient background of susceptibility to DLI. Sorafenib is mainly metabolized in the liver, but

impaired metabolism of sorafenib because of decreased hepatic function reserve has not been proven according to the result that no clinically relevant difference was observed in pharmacokinetics between patients of Child–Pugh class A and class B in the Phase 1 study of sorafenib in Japanese patients with HCC [23].

Deteriorated performance status, pre-existing chronic fibrosing ILD, and smoking history were factors that contributed to DLI in NSCLC patients treated with gefitinib [9, 10, 20] and erlotinib [12]. Similarly, deteriorated systemic condition in advanced HCC patients may be attributed to decreased hepatic function reserve and/or multiple metastases. In addition, several reports suggest pathogenic role of chronic hepatitis C viral infection and its drug treatment in DLI, although the association between them remains controversial [24]. It is interesting that the difference in patient background because of underlying tumor types may be responsible for the occurrence of DLI.

In the present analysis, 18 out of 34 patients who had imaging data available showed DAD pattern. Furthermore, it is important to keep in mind that two-third of the patients with DAD pattern had a fatal outcome.

Diffuse alveolar damage is characterized histologically by the presence of alveolar airspace and interstitial oedema, hyaline membrane formation, and proliferation of type 2 pneumocytes [16, 19]. It manifests radiographically as bilateral hetero- or homogeneous opacities, usually in the mid and lower lungs, and on high-resolution CT scans as scattered or diffuse areas of ground-glass opacity, and architectural distortion can occur [18, 19]. The image findings from individual patients showed that several patients eventually presented the DAD pattern, although their image findings at onset were faint ground-glass opacity. The importance of early recognition, close observation, and initiation of treatment for DLI at an early stage should be emphasized to physicians. In addition, it is important to consult with specialists for pulmonary medicine at an early stage.

Currently, specific guidelines for the treatment of DLI are not available, and treatment tends to be administered on an empirical basis. High-dose methylprednisolone for several days followed by tapering of the dose is commonly used, in addition to withdrawal of the suspected drug [7, 25]. Overall, the treatment for DLI is considered appropriate in the present analysis. It should be noted, however, that the DLI of radiological DAD pattern often leads to poor prognosis despite early recognition and early initiation of treatment.

Limitations of the present analysis include possible bias in patients with availability of detailed information, and lack of statistical analysis specifying prognostic factors or risk factors for sorafenib-associated lung injury. The observation period was not uniformly determined in the present analysis, because of an observational post-marketing surveillance setting. In addition, we did not perform in-depth analyses of the levels of serum markers, arterial blood gas, or the results of bronchoalveolar lavage/transbronchial lung biopsy. The information of pathological examination to confirm the image evaluation is also limited.

The present analysis will provide useful information about DLI to health care professionals involved in the treatment using sorafenib. An analysis to specify risk factors will be reported in the final result of SDUI. Further investigations are required to determine the difference in DLI according to causative drugs, cancer types, and ethnicity.

Acknowledgments We would like to express our deepest appreciation to the medical doctors participating in the Nexavar® SDUI in RCC and HCC in Japan as well as the reporting physicians for DLI patients for providing detailed information.

Conflict of interest Y. H-Y. is an employee of Bayer Yakuhin, Ltd. A.G. has received consulting fees from Bayer Yakuhin, Ltd. as a member of the ILD Ad-board in the subject of this manuscript. H.T. has received consulting fees from Bayer Yakuhin, Ltd. as a member of the ILD Ad-board in the subject of this manuscript. Y.I. has received consulting fees from Bayer Yakuhin, Ltd. as a member of the ILD Ad-board in the subject of this manuscript. F.S. has received consulting fees from Bayer Yakuhin, Ltd. as a member of the ILD Ad-board in the subject of this manuscript. T.J. has received consulting fees from Bayer Yakuhin, Ltd. as a member of the ILD Ad-board in the subject of this manuscript, those from Chugai Seiyaku, Ltd., as a member of the ILD Ad-board, and conducted honorary lectures with support from Daiichi-Sankyo Seiyaku, Ltd., Kyorin Seiyaku, Ltd., and Eizai Ltd. K.F. has received consulting fees from Bayer Yakuhin, Ltd. as a member of the ILD Ad-board in the subject of this manuscript. S.K. has received consulting fees from Bayer Yakuhin, Ltd. as a member of the ILD Ad-board in the subject of this manuscript.

References

1. Wilhelm SM, Carter C, Tang L et al (2004) BAY 43-9006 exhibits broad spectrum oral antitumor activity and targets the RAF/MEK/ERK pathway and receptor tyrosine kinases involved in tumor progression and angiogenesis. Cancer Res 64:7099–7109
2. Chang YS, Adnane J, Trail PA et al (2007) Sorafenib (BAY 43-9006) inhibits tumor growth and vascularization and induces tumor apoptosis and hypoxia in RCC xenograft models. Cancer Chemother Pharmacol 59(5):561–574
3. Liu L, Cao Y, Chen C et al (2006) Sorafenib blocks the RAF/MEK/ERK pathway, inhibits tumor angiogenesis, and induces tumor cell apoptosis in hepatocellular carcinoma model PLC/PRF/5. Cancer Res 66(24):11851–11858
4. Escudier B, Eisen T, Stadler WM et al (2007) Sorafenib in advanced clear-cell renal-cell carcinoma. N Engl J Med 356(2):125–134
5. Llovet JM, Ricci S, Mazzaferro V et al (2008) Sorafenib in advanced hepatocellular carcinoma. N Engl J Med 359:378–390
6. Bayer Yakuhin, Ltd. (2008) Safety information for acute lung injury/interstitial pneumonia for Nexavar® 200 mg tablets (in Japanese)

7. Müller NL, White DA, Jiang H et al (2004) Diagnosis and management of drug-associated interstitial lung disease. Br J Cancer 91(Suppl 2):S24–S30
8. Yoshii N, Suzuki T, Nagashima M et al (2011) Clarification of clinical features of interstitial lung disease induced by irinotecan based on postmarketing surveillance data and spontaneous reports. Anticancer Drugs 22(6):563–568
9. Gemma A (2009) Drug-induced interstitial lung diseases associated with molecular-targeted anticancer agents. J Nihon Med Sch 76(1):4–8
10. Ministry of Health, Labor and Welfare (2004) The report of prospective study for Iressa® tablets 250. Pharmaceuticals and Medical Devices Safety Information No. 206 (in Japanese)
11. Ando M, Okamoto I, Yamamoto N et al (2006) Predictive factors for interstitial lung disease, antitumor response, and survival in non-small-cell lung cancer patients treated with gefitinib. J Clin Oncol 24(16):2549–2556
12. Hotta K, Kiura K, Takigawa N et al (2010) Comparison of the incidence and pattern of interstitial lung disease during erlotinib and gefitinib treatment in Japanese patients with non-small cell lung cancer: the Okayama Lung Cancer Study Group experience. J Thorac Oncol 5(2):179–184
13. Ide S, Soda H, Hakariya T et al (2010) Interstitial pneumonia probably associated with sorafenib treatment: an alert of an adverse event. Lung Cancer 67(2):248–250
14. Bayer Yakuhin, Ltd. (2011) Nexavar® Tablets 200 mg, Third Interim Report of Special Drug Use Investigation for unresectable or metastatic renal cell carcinoma (in Japanese)
15. Bayer Yakuhin, Ltd. (2010) Nexavar® Tablets 200 mg, Interim Report of Special Drug Use Investigation for unresectable hepatocellular carcinoma (in Japanese)
16. Rossi SE, Erasmus JJ, McAdams HP et al (2000) Pulmonary drug toxicity: radiologic and pathologic manifestations. Radiographics 20:1245–1259
17. Akira M, Ishikawa H, Yamamoto S (2002) Drug-induced pneumonitis: thin-section CT findings in 60 patients. Radiology 224(3):852–860
18. Ichikado K, Suga M, Gushima Y et al (2000) Hyperoxia-induced diffuse alveolar damage in pigs: correlation between thin-section CT and histopathologic findings. Radiology 216(2):531–538
19. Cleverley JR, Screaton NJ, Hiorns MP et al (2002) Drug-induced lung disease: high-resolution CT and histological findings. Clin Radiol 57:292–299
20. Kudoh S, Kato H, Nishiwaki Y et al (2008) Interstitial lung disease in Japanese patients with lung cancer: a cohort and nested case-control study. Am J Respir Crit Care Med 177(12):1348–1357
21. Cohen MH, Williams GA, Sridhara R et al (2004) United States Food and Drug Administration Drug Approval summary: gefitinib (ZD1839; Iressa) tablets. Clin Cancer Res 10:1212–1218
22. Nyberg F, Barratt BJ, Mushiroda T et al (2011) Interstitial lung disease in gefitinib-treated Japanese patients with non-small-cell lung cancer: genome-wide analysis of genetic data. Pharmacogenomics 12(7):965–975
23. Furuse J, Ishii H, Nakachi K et al (2008) Phase I study of sorafenib in Japanese patients with hepatocellular carcinoma. Cancer Sci 99(1):159–165
24. Moorman J, Saad M, Kosseifi S et al (2005) Hepatitis C virus and the lung: implications for therapy. Chest 128(4):2882–2892
25. American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the idiopathic interstitial pneumonias (2002). Am J Respir Crit Care Med 165(2):277–304

Inflammation, Vol. 36, No. 2, April 2013 (© 2012)
DOI: 10.1007/s10753-012-9545-3

Mycoplasma pneumoniae Extract Induces an IL-17-Associated Inflammatory Reaction in Murine Lung: Implication for Mycoplasmal Pneumonia

Daisuke Kurai,[1] Kazuhide Nakagaki,[2] Hiroo Wada,[1] Takeshi Saraya,[1] Shigeru Kamiya,[3] Yasunori Fujioka,[4] Koh Nakata,[5] Hajime Takizawa,[1] and Hajime Goto[1,6]

Abstract—*Mycoplasma pneumoniae* (Mp) may cause immune cell reactions as pivotal aspects of this clinically common respiratory pathogen. Our aim is to determine if Mp extract induces a cellular immune response associated with interleukin (IL)-17, leading to lung inflammation and lung injury. BALB/c mice were immunized with Mp extract intraperitoneally followed by its intratracheal administration, to mimic repeated Mp infection found in humans (repeated inoculation, RI group). Those with a single inoculation were compared as single inoculation group (SI group). Analysis of bronchoalveolar lavage fluid (BALF) demonstrated that keratinocyte-derived cytokine, tumor necrosis factor-α, and IL-6 were produced and peaked on days 0.5 or 1, followed by IL-17 on day 2. Levels of these mediators in BALF were higher in RI group than SI group ($P<0.05$). Further, significantly more neutrophils were recruited to the lungs of the RI group ($P<0.05$). These observations suggest that IL-17 is involved in the prolonged induction of neutrophils in mice treated with Mp extract.

KEY WORDS: *Mycoplasma pneumoniae*; IL-17; IL-23; neutrophil.

INTRODUCTION

Mycoplasma pneumoniae (Mp) is a well-known cause of community-acquired pneumonia that can induce a cellular immune response, leading to inflammation and lung injury [1]. Even cases of pulmonary diseases caused by macrolide-resistant Mp have been successfully treated with macrolide, an agent that has an immunomodulatory effect [2], and cellular host defense reactions are enhanced by repeated Mp stimulation in animal models [3]. The clinical manifestation of Mp infection in humans varies with age, as younger children infected with Mp do not tend to develop pneumonia, whereas older children and adults do tend to develop pneumonia [4–7].

The symptoms of Mp pneumonia are more severe than expected in individuals who have been immunized with Mp vaccine [8]. Studies have demonstrated that interleukin (IL)-4 in bronchoalveolar lavage fluid (BALF) and IL-5 in blood were elevated in patients with Mp pneumonia [9, 10]. Several reports further demonstrated that corticosteroid was clinically beneficial for patients with Mp pneumonia, suggesting that anti-immune and anti-inflammatory actions of corticosteroid were effective against these cellular host defense reactions [11, 12]. Taken together, it has been speculated that cellular immune reactions play an important role in induction of lung inflammation and lung injury by Mp.

In order to mimic the effect of Mp on such immune cell reactions, we previously established a murine model

[1] Department of Respiratory Medicine, Kyorin University School of Medicine, 6-20-2 Shinkawa Mitaka-shi, Tokyo 181-8611, Japan
[2] College of Veterinary Medicine, Nippon Veterinary and Life Science University, 1-7-1, Kyonan-cho Musashino-shi, Tokyo 180-8602, Japan
[3] Department of Infectious Diseases, Kyorin University School of Medicine, 6-20-2 Shinkawa Mitaka-shi, Tokyo 181-8611, Japan
[4] Department of Pathology, Kyorin University School of Medicine, 6-20-2 Shinkawa Mitaka-shi, Tokyo 181-8611, Japan
[5] Bioscience Medical Research Center, Niigata University Medical and Dental Hospital, 1-757 Asahimachi-Dohri, Chuo-ku, Niigata 951-8520, Japan
[6] To whom correspondence should be addressed at Department of Respiratory Medicine, Kyorin University School of Medicine, 6-20-2 Shinkawa Mitaka-shi, Tokyo 181-8611, Japan. E-mail: h510@ks.kyorin-u.ac.jp

of Mp pneumonia for which we employed Mp extract, rather than live Mp, in order to avoid the infectious aspect of the disease [13, 14]. Histological analysis of lung tissues of this mouse model, which had an immune cell reaction against Mp, showed lymphoplasmacytic inflammation in the peri-bronchial area, which is the same pathology as that observed in humans suffering from severe Mp pneumonia [14]. Interestingly, comprehensive analysis of inflammatory mediators in these mice demonstrated that IL-17 levels increased in BALF following Mp inoculation [13]. IL-17 concentration and neutrophil counts in BALF were elevated in parallel when mice were inoculated with live Mp [15]. Similar IL-17-linked signal activation was observed in patients suffering from Mp pneumonia who showed significantly higher levels of serum IL-17 than patients with streptococcal pneumonia [16]. Since it has been reported that IL-17 plays a role in the transition from innate immunity to adaptive immunity [17, 18], these observations led to the hypothesis that some components of the Mp extract induce IL-17, which then results in excessive inflammatory cell reactions in Mp pneumonia.

The mouse models of Mp pneumonia that we previously established [13, 14] were pretreated with alum-adjuvant, which may affect the expression of IL-17 and associated molecules. In the present study, a novel mouse model of Mp pneumonia that did not employ alum-adjuvant treatment was prepared. Furthermore, BALB/c mice were immunized with Mp extract intraperitoneally followed by its intratracheal administration, to mimic repeated Mp infection found in humans (repeated inoculation, RI group). Those with a single inoculation were compared as single inoculation group (SI group) in order to clarify whether IL-17 levels are enhanced by injection of Mp extract.

MATERIALS AND METHODS

Preparation of Mp Extract

Mp extract was prepared from cultured *M. pneumoniae* (ATCC 29342) in a pleuropneumonia-like organism liquid broth containing 20 % horse serum and 10 units/mL penicillin G (Nikkenkagaku, Kyoto, Japan), according to a previously published method with some modification [13]. In brief, cultured Mp was centrifuged and repeatedly washed with Hanks' balanced salt solution (HBSS) (Invitrogen, Grand Island, NY). The sediment was suspended in HBSS and sonicated. This suspension was centrifuged and the supernatant was defined as the Mp extract. The amount of Mp extract was defined in terms of protein concentration, and the Mp extract was prepared in HBSS at a concentration of 1.0 μg/μL.

Primary Culture of Lung-Derived Cells

Cells were prepared from lung tissues originating from four 12-week-old BALB/c female mice (Charles River Laboratories, Kanagawa, Japan), which were anesthetized and then sacrificed after their blood was replaced with saline. Their lungs were removed and cut into pieces, and the tissues were incubated for 2 h at room temperature in Dulbecco's modified Eagle's medium (DMEM) (Invitrogen, Grand Island, NY, USA) that contained 300 U/mL collagenase type 1 (Worthington Biochemical Corporation, NJ, USA), 0.01 % gentamicin (GM), and 0.1 % amphotericin-B (AMPH-B). The suspended cells were then harvested, washed with DMEM containing 0.01 % GM, 0.1 % AMPH-B, and 10 % fetal calf serum (FCS, Invitrogen), passed five times through 40-μm nylon mesh cell strainers (BD FALCON, Franklin Lakes, NJ, USA), seeded at a concentration of 5.0×10^4 cells/well in a 96-well plate, and incubated at 37 °C in 5 % CO_2. After 24 h, the culture media were replaced with DMEM containing 0.01 % GM, 0.1 % AMPH-B, 10 % FCS, and 100 U/mL polymyxin B. To assess the time-dependent release of cytokines, the cells were treated with or without Mp extract (50 μg/mL). After incubation of the cells for 24, 48, and 72 h with the Mp extract, the media were collected for measurement of IL-23 and IL-17 levels. Media were also collected after 72 h incubation with different concentrations of the Mp extract (0, 1.9, 5.6, 16.7, and 50 μg/mL) to evaluate dose-dependent stimulating effects of Mp extract. Each assay was performed in quadruplicates.

Mp Inoculation and Sampling

The protocol of Mp inoculation and sampling is shown in Fig. 1. Twelve-week-old BALB/c female mice (Charles River Laboratories, Kanagawa, Japan) were assigned to one of two groups. Mice in the SI group were intratracheally inoculated with 50 μg/50 μL of Mp extract alone on day 0. Mice in the RI group were intraperitoneally injected twice with 50 μg/250 μL of Mp extract 28 and 21 days before intratracheal injection with 50 μg/50 μL of the Mp extract (day 0). In RI group, the Mp extract was administered intraperitoneally

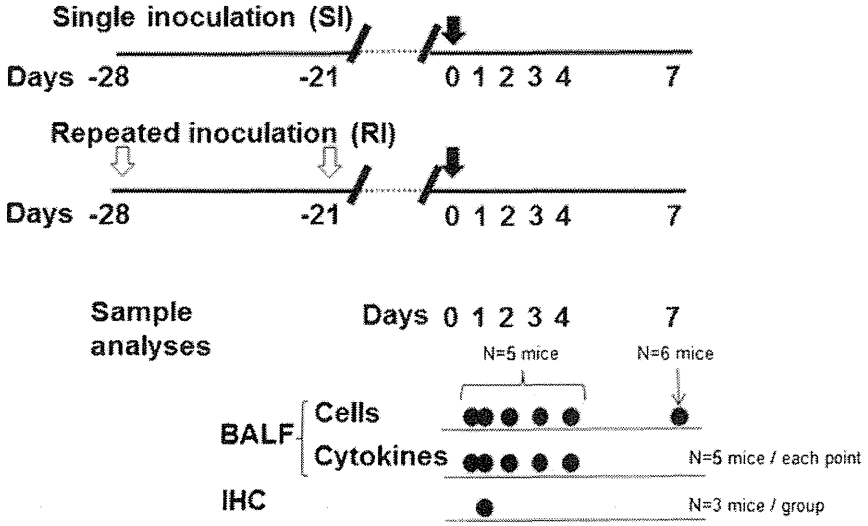

Fig. 1. Mp extract inoculation and sampling protocol. Mice that were SI were administered with Mp extract intratracheally on day 0. Mice that were RI were administered with Mp extract intraperitoneally on days −28 and −21 and were then inoculated intratracheally with the Mp extract on day 0. BALF was obtained on days 0.5, 1, 2, 3, 4, and 7 for analysis of cells (n=5 per each on days 0.5, 1, 2, 3, and 4; n=6 on day 7) and cytokines (n=5 per each point). Lungs were obtained on day 1 for IHC analysis (n=3 per each group).

in order to avoid damaging the trachea by repeated inoculations and to obtain reproducible results. Mice were anesthetized with pentobarbital and lavaged twice with 1 mL of HBSS through a catheter *via* the trachea. BALF was subsequently collected and was centrifuged at 400×g for 5 min. Total and differential cell counts were calculated for BALF (number of mice tested was indicated in each figure) collected on days 0.5, 1, 2, 3, 4, and 7. Cytokine levels were measured in BALF collected on days 0.5, 1, 2, 3, and 4 (number of mice tested was indicated in each figure). In preliminary experiments, any of cytokines tested were not detected in BALF collected on day 7. Immunohistochemical analysis of IL-23 expression in lungs (number of mice tested was indicated in each figure) was conducted on day 1 because preliminary experiments demonstrated that the percentage of IL-23-positive cells in lung tissue was higher on day 1 than on day 0.5. All animal experiments were performed in accordance with the Institutional Animal Care and Research Advisory Committee at Kyorin University.

Cytokine Measurements

The levels of IL-17, keratinocyte-derived cytokine (KC), tumor necrosis factor-α (TNF-α), IL-6, interferon-γ (IFN-γ), IL-4, and IL-12 in BALF were measured using a protein multiplex immunoassay kit (Biosource International, Camarillo, CA, USA) and a multiplex bead array (Luminex 200, Luminex, Austin, TX, USA), according to the manufacturer's instructions. The concentrations of IL-17 and IL-23 in the primary cell cultures were measured using an enzyme-linked immunosorbent assay kit (Quantikine mouse IL-17, Quantikine mouse IL-23, R&D Systems, Minneapolis, MN, USA).

IL-23 Immunohistochemistry

Mouse lungs were removed on day 1, fixed in 4 % paraformaldehyde, embedded in paraffin, and sectioned. After deparaffinization, the specimens were stained with an anti-IL-23 rabbit polyclonal antibody to IL-23p19 (Abcam, Cambridge, UK) at a dilution of 1:100, followed by staining with a peroxidase-conjugated secondary antibody at a dilution of 1:100. The number of IL-23-positive and IL-23-negative cells in ten microscope fields (330×430 μm) per mouse (N=3) was counted in a blinded fashion, and the percentage of the IL-23-positive cells to the total cells was calculated.

Statistical Analysis

All data are expressed as means ± standard deviation. Statistic analysis was performed using the Statistical Package for Social Sciences (SPSS) software (SPSS, Chicago, IL, USA). The Kruskal–Wallis test was

used to evaluate variance among all groups. If a significant variance was found, the Mann–Whitney test was used to determine significant differences between individual groups. $P<0.05$ was considered to represent a statistically significant difference.

RESULTS

Secretion of IL-23 and IL-17 by Primary Cultured Cells

Incubation of mouse lung-derived primary cell cultures with Mp extract (50.0 μg/mL) resulted in an increase in the concentration of IL-23 and IL-17 in the culture media when compared with control (Mp extract 0 μg/mL). The IL-23 concentration peaked at 24 h, whereas the IL-17 concentration peaked at 72 h (Fig. 2a). The IL-23 concentration was significantly higher at 24 and 48 h than that at 72 h. The IL-17 concentration at 48 and 72 h was significantly higher than that at 24 h.

Incubation of the lung-derived primary cultures with increasing concentrations of Mp extract (from 0 to 50 μg/mL) for 72 h resulted in a dose-dependent increase in IL-23 concentration in the media, which peaked at Mp extract concentration of 5.6 μg/mL (Fig. 2b). The IL-17 concentration in the media also significantly increased at an Mp extract concentration of no less than 5.6 μg/mL; to show a concentration dependency up to the highest Mp concentration of 50 μg/mL (Fig. 2c).

Analysis of the Cytokines in BALF

We measured the level of IL-17 and its associated cytokines in the BALF collected from mouse groups that were singly or repeatedly inoculated. IL-17 was detected

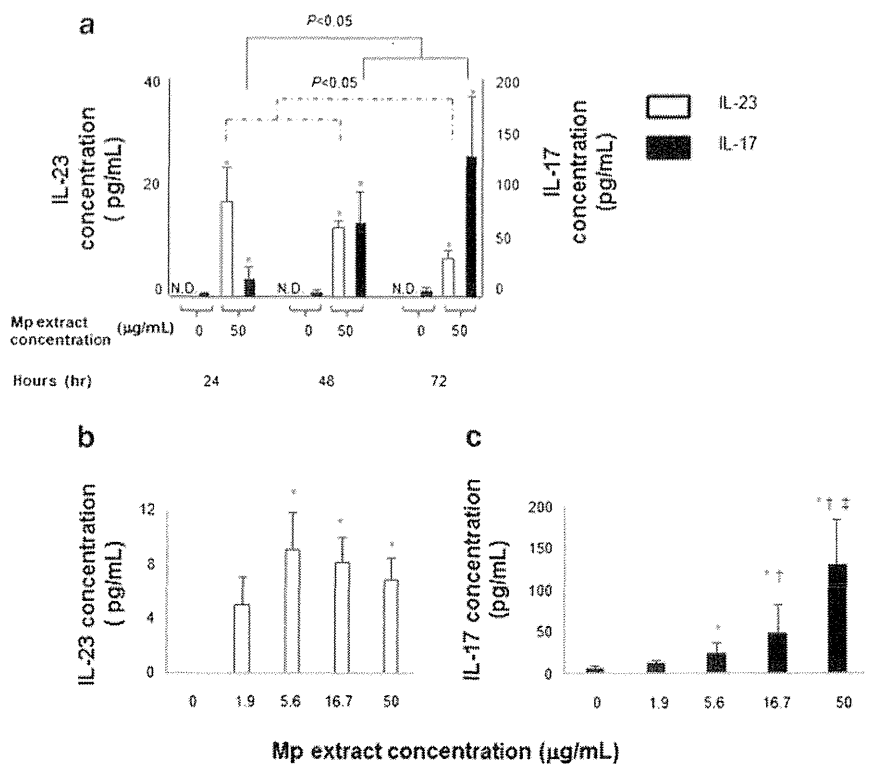

Fig. 2. Effect of incubation of primary lung cultures with Mp extract on the concentration of IL-23 and IL-17 in the culture media. **a** The concentration of IL-23 (*open bars*) and IL-17 (*solid bars*) in primary lung culture media was measured at 24, 48, and 72 h after the addition of *M. pneumonia* (*Mp*) extract at a final protein concentration of 0 and 50.0 μg/mL. The concentration of IL-23 (**b**) and IL-17 (c) in primary lung culture media was measured at 72 h after addition of the indicated concentration of the Mp extract. Significant differences are indicated by the following symbols: $*P<0.05$ compared with no added Mp extract (0 μg/mL); $^{\dagger}P<0.05$ compared with the 1.9-μg/mL Mp extract; and $^{\ddagger}P<0.05$ compared with the 5.6-μg/mL Mp extract. Data are expressed as means ± standard deviation. Each assay was performed in quadruplicates. *N.D.* indicates "not detected."