Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Trapnell BC, Inoue Y, Bonella F, et al. Inhaled molgramostim therapy in autoimmune pulmonary alveolar proteinosis. N Engl J Med. DOI: 10.1056/NEJMoa1913590

(PDF updated September 8, 2020)

Table of Contents **Page**

SECTION 1. Complete List of Impala Trial Primary Investigators and Institutions

Investigators

Ali Ataya, MD[1], Antonio Morais, MD[2], Bruce C. Trapnell, MD[3], Bryan D. Kraft, MD[4], Cliff Morgan, BM, FRCA, FFICM[5], Daniel A. Culver, MD[6], Elisabeth Bendstrup, MD, PhD[7]. Erdogan Cetinkaya, MD[8], Etsuro Yamaguchi, MD, PhD[9], Francesco Bonella, MD, PhD[10], Francesca Mariani, MD[11], Jin Woo Song, MD PhD[12], John Wheatley, MD[13], Jürgen Behr, MD[14], Koh Nakata, MD PhD[15], Lauren Troy, MD[16,] Leonardo Ferreira, MD[17], Makoto Kobayashi, MD[18]. Marcel Veltkamp, MD PhD[19], Maria Molina, MD[20], Michael Kreuter, MD[21], Mikhail M. Ilkovich, MD[22], Mordechai R. Kramer, MD[23], Pan Pyo Chung, MD PhD[24], Robert Bals, MD[25], Robert Slivka, MD PhD[26], Romain Lazor, MD[27], Daniel Droemann, MD[28], Stéphane Jouneau, MD, PhD[29], Sun Mi Choi, MD[30], Spyros A. Papiris, MD[31], Tisha Wang, MD[32], Tomohisa Baba, MD[33], Yoshikazu Inoue, MD, PhD[34]

Institutions

[1]University of Florida, Division of Pulmonary Critical Care, Gainesville, FL, USA; [2]Hospital Sao Joao, Serviço de Pneumologia, Porto, Portugal; [3]Translational Pulmonary Science Center, Cincinnati Children's Hospital Medical Center, Cincinnati, OH; [4]Duke University Medical Center, Division of Pulmonary, Allergy and Critical Care Medicine, Durham, NC, USA; [5]Department of Intensive Care Unit, Royal Brompton Hospital, London, England; [6]Cleveland Clinic, Department of Pulmonary, Allergy and Critical Care, Cleveland, OH, USA; [7]Department of Respiratory Diseases & Allergy, Aarhus University Hospital, Aarhus, Denmark, [8]University of Health Sciences Turkey, Yedikule Chest Diseases and Thoracic Surgery Education and Research Hospital, Istanbul, Turkey; [9]Aichi Medical University Hospital, Nagakute, Aichi, Japan; [10] Outpatients Clinic for Interstitial and Rare Lung Disease, Ruhrlandklinik University Hospital, Essen, Germany; [11]Pneumology Unit, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy; [12]Asan Medical Center, Division of Pulmonary and Critical Care Medicine, University of Ulsan College of Medicine, Seoul, South Korea; [13]Westmead Hospital, Department of Respiratory Medicine, Sydney, Australia; [14]Asklepios Fachkliniken München-Gauting, Klinik für Pneumologie, Germany; [15]Niigata University Medical and Dental Hospital Bioscience Medical Research Center, Japan; [16]Royal Prince Alfred Hospital, Sydney, Australia; [17]CHLN, Hospital Pulido Valente, Serviço de Pneumologia, Hospital de Dia de Pneumologia, Lisbon, Portugal; [18]Division of Respiratory medicine, Tohoku University Hospital, Myaigi, Japan; [19] ILD Center of Excellence, Department of Pulmonology, St. Antonius Hospital, Nieuwegein, The Netherlands; [20]Hospital University de Bellvitge, Servei de Pneumologia, Barcelona, Spain; [21] Center for Interstitial and Rare Lung Diseases,Pulmonology, Thoraxklinik, Heidelberg University Hospital, Heidelberg, and German Center for Lung Research, Heidelberg, Germany; [22]Pulmonary Clinic of St. Petersburg Pavlov State Medical University, St. Petersburg, Russia; [23]Institute of Pulmonary and Allergy Medicine, Rabin Medical Center, Petah Tikva, Israel; [24]Samsung Medical Center, Division of Pulmonary and Critical Care Medicine, Department of Medicine, Sungkyunkwan University School of Medicine, Seoul, South Korea; [25]Universitätsklinikum des Saarlandes Klinik für Innere Medizin V Pneumologie, Allergologie, Beatmungs- und Umweltmedizin Gebäude 91, Homburg, Germany; [26]Pulmonary Department, National Institut for TB, Lung Diseases and Chest Surgery, Vysne Hagy, Slovakia; [27]Centre Hospitalier Universitaire Vaudois, Service de Pneumologie, Lausanne, Switzerland; [28]Universitätsklinikum Schleswig-Holstein, Zentralklinikum Lübeck, Medizinische Klinik III – Pneumologie, Lübeck, Germany; [29]Respiratory Diseases Department, Pontchaillou Hospital, IRSET UMR 1085, Rennes 1 University, Rennes, France; [30]Seoul National University Hospital, Seoul, South Korea; [31]2nd Pulmonary Medicine Department, General University Hospital "Attikon", Medical School, National and Kapodistrian University of Athens, Greece; [32]UCLA, Pulmonary and Critical Care, David Geffen School of Medicine, Los Angeles, CA, USA; [33]Kanagawa Cardiovascular and Respiratory Center, Yokohama, Japan; [34]National Hospital Organization Kinki-Chuo Chest Medical Center, Osaka, Japan.

SECTION 2. Detailed Inclusion and Exclusion Criteria

Inclusion Criteria:

- aPAP diagnosed by CT, or by biopsy, or by bronchoalveolar lavage (BAL), and by increased GM-CSF autoantibodies in serum
- Stable or progressive aPAP (i.e. absolute VC not improved by more than 5% and/or DLCO not improved by more than 10% - assessed from medical records) during a minimum period of two months prior to the Baseline visit
- PaO_2 <75 mmHg/<10 kPa at rest, OR desaturation of >4 percentage points on the distance covered in a 6-Minute Walk Test (6MWT)
- An A-aDO_2 of minimum 25 mmHg/3.33 kPa
- Female or male ≥18 years of age
- Females who have been post-menopausal for >1 year or females of childbearing potential after a confirmed menstrual period using a highly efficient method of contraception (i.e. a method with <1% failure rate such as combined hormonal contraception, progesterone-only hormonal contraception, intrauterine device, intrauterine hormone-releasing system, bilateral tubal occlusion, vasectomised partner, sexual abstinence), during and until 30 days after last dose of trial treatment. Females of childbearing potential must have a negative serum pregnancy test at Screening (Visit 1) and a negative urine pregnancy test at dosing at Baseline (Visit 2) and must not be lactating
- Males agreeing to use condoms during and until 30 days after last dose of trial treatment, or males having a female partner who is using adequate contraception as described above
- Willing and able to provide signed informed consent
- Willing and able to comply with scheduled visits, treatment plan, laboratory tests, and other trial procedures specified in the protocol as judged by the investigator

Exclusion Criteria:

- Diagnosis of hereditary or secondary pulmonary alveolar proteinosis (PAP)
- WLL within one month of Baseline
- Treatment with GM-CSF within three months of Baseline
- Treatment with rituximab within six months of Baseline
- Treatment with plasmapheresis within three months of Baseline
- Treatment with any investigational medicinal product within four weeks of Screening
- Concomitant use of sputum modifying drugs such as carbocystein or ambroxol
- History of allergic reactions to GM-CSF
- Connective tissue disease, inflammatory bowel disease or other autoimmune disorder requiring treatment associated with significant immunosuppression, e.g. more than 10 mg/day systemic prednisolone
- Previous experience of severe and unexplained side-effects during aerosol delivery of any kind of medicinal product
- History of, or present, myeloproliferative disease or leukaemia
- Known active infection (viral, bacterial, fungal or mycobacterial)
- Apparent pre-existing concurrent pulmonary fibrosis
- Any other serious medical condition which in the opinion of the investigator would make the subject unsuitable for the trial

SECTION 3. List of Outcome Measures

Primary and Key Secondary Efficacy End Points

Primary Efficacy End Point:

- Absolute change from baseline of A-aDO_2 after 24-weeks treatment

Key Secondary Efficacy End Points:

- Change from baseline in 6MWD after 24-weeks treatment
- Change from baseline in SGRQ total score after 24-weeks treatment
- Time to WLL during 24-weeks treatment

Further Secondary and Exploratory Efficacy Variables

Double-blind Treatment Period

- Number of WLL procedures performed
- DLCO (% predicted), FEV_1 (% predicted), FVC (% predicted), VC (% predicted), PaO_2
- Tolerance to exercise (improvement in 6MWT-Distance or desaturation)
- Dyspnea and cough scores
- CT score

Exploratory Efficacy Variables

- SGRQ 4-point responders
- EQ-5D-5L responders
- Serum biomarkers: Lactate Dehydrogenase (LDH), Krebs von den Lungen-6 (KL-6), Carcinoembryonic antigen (CEA), Cytokeratin 19 Fragment (Cyfra 21-1), Surfactant Protein A (SP-A), Surfactant Protein B (SP-B), Surfactant Protein C (SP-C), Surfactant Protein D (SP-D)
- Anti-GM-CSF antibody titers
- Serum levels of GM-CSF
- Need for oxygen supplement therapy
- Distribution of Disease Severity Score

Follow-up Period

- Number of patients requiring WLL or other treatment for PAP
- Time to WLL
- A-aDO_2 , FEV_1 (% predicted), FVC (% predicted), VC (% predicted), DLCO (% predicted), PaO_2
- Anti-GM-CSF antibody titers
- Distribution of Disease Severity Score

Post-hoc Efficacy Variables

- SGRQ 8-, and 12-point responders
- SGRQ Component scores
- SGRQ Total Score during the follow-up period

Safety Outcomes:

- Number of AEs, SAEs, ADRs, severe AEs and AEs leading to treatment discontinuation, including clinically significant changes in laboratory tests and electrocardiogram (ECG) variables, during double-blind and follow-up treatment periods
- Physical examination
- Electrocardiograms
- Vital signs
- Laboratory safety assessments (complete blood count, liver function tests, etc.)

SECTION 4. Supplemental Methods

Clinical Management of the Study Patients

For ethical reasons, all enrolled patients were permitted to receive whole lung lavage (WLL) to manage progression of aPAP lung disease or oxygen to maintain adequate blood oxygen levels if required.

Whole Lung Lavage Therapy

WLL was permitted at any time during the trial as rescue therapy for progression of aPAP as required based on the clinical judgement of the clinical site investigator. If the patient received WLL, the following information was to be documented in the electronic case report forms: the date when the decision to perform WLL was made.

Supplemental Oxygen Therapy

Use of supplemental oxygen therapy was permitted at any time during the trial according to the clinical site investigator's judgement. When used, the following information was to be documented in concomitant medication/therapy page of the electronic case report forms: the date(s) of usage, the administration method (e.g., nasal canula), and the rate of administration (e.g., the flow rate in liters per minute).

Assignment to Treatment Groups and Randomization of Study Patients

Each patient was assigned a unique, site-specific study participant identification number generated automatically by the interactive web response system (IWRS) maintained by S-Clinica; the unique patient identification numbers were used throughout the trial. At Baseline (Visit 2), eligible patients were randomized centrally assigned to one of the three Double-blind treatment groups automatically by the IWRS. Randomisation was stratified according to receipt of a WLL within 2 months prior to Baseline (Visit 2). Medication kits administered during the Open-Label treatment period were not blinded.

Study Drugs

The active pharmaceutical ingredient (API) consisted of molgramostim (ATC code: L03AA03) produced by using recombinant DNA technology in an *Escherichia coli* expression system, which results in production of non-glycosylated recombinant human GM-CSF. Structurally, molgramostim is a compact globular protein containing a four-helix bundle and a closely packed hydrophobic core. X-ray crystal structure of molgramostim indicate α-helical and β-sheet content of 40-50% and 6-10%, respectively.[1] The drug product is the drug substance formulated by dilution with excipients in a citrate phosphate buffer at pH 7.0 to 7.4 contained in a glass vial. Each vial of molgramostim contains 250 µg/mL molgramostim in 1.2 mL solution. The matching placebo consists of a citrate phosphate buffer at pH 7.0 to 7.4 containing the same excipients but without the active molgramostim. All manufacturing and packaging is performed in accordance with current Good Manufacturing Practice (GMP).

Study Drug Kits

Individual medication kits containing double-blind trial medication was supplied in adequate amounts at each dispensing visit. Every kit was labelled with a unique medication number assigned automatically by the IWRS system, which also appeared on a tear-off part of the label which was applied to a subject diary used to document treatment compliance.

Delivery Device

The PARI eFlow nebuliser system (PARI Pharma GmbH, Germany) was used to administer the investigational agent (IA). The eFlow Nebulizer Handset is a single patient use, reusable electronic

nebulizer. It includes a fine particle aerosol generator (perforated vibrating membrane) defined by a 30L mesh and an aerosol chamber that can produce aerosols with high density of active drug, precisely defined droplet size and a high proportion of respirable droplets.

Training

All patients, investigators and trial nurses were trained in administration of the AI and maintenance of the nebulizer device. The training of the patients was arranged prior to administration of the patient's first dose of IA and checked in clinic on first dosing. The patient also received written instructions. The patient administered the first dose of IA post randomisation at the Baseline visit (Visit 2) under the supervision of trial personnel.

Molgramostim/GM-CSF autoantibody Evaluations

Quantification of anti-molgramostim/human GM-CSF antibodies in patient serum

The potential development of anti-drug antibodies directed against molgramostim (recombinant human GM-CSF) was evaluated by measuring the titer of anti-GM-CSF antibodies in serum from study patients using a homogeneous electrochemiluminescence-based bridging assay developed by the sponsor. Briefly, molgramostim was labeled with biotin and mixed with serum to permit formation of immune complexes between the biotinylated molgramostim and anti-GM-CSF antibodies present in serum. Immune complexes were captured on streptavidin-coated assay plates and antibodies were detected using Ru^{2+}-containing Sulfo-tag™. Importantly, since molgramostim and endogenous GM-CSF have the same amino acid sequence, antibodies against these two molecules are both detected by this assay and cannot be distinguished. The methods were validated prior to analysis of study samples. In each run, high concentrated controls, low concentrated controls and negative controls were included. During the validation a plate-specific floating screening cut point was determined. Samples that were above the cut point in a screening assay were confirmed positive by inhibiting the signal with high levels of rhGM-CSF. Samples showing over 40.1% inhibition in the confirmatory assay were scored positive and semi-quantified by titration by analysis in serial 1:2 dilution(s). The highest titer equal to or above the floating cut point was reported.

Quantification of the GM-CSF-neutralizing capacity of patient-derived serum

The GM-CSF neutralizing capacity of patient serum (caused by GM-CSF autoantibodies and or anti-molgramostim antibodies) was measured using a cell proliferation assay using cultured TF-1 cells (DSMZ no. ACC 334), which normally proliferate in response to stimulation by GM-CSF.[3] In this assay, neutralizing antibodies reduce GM-CSF-mediated growth stimulation.[4] Briefly, cell growth was measured by spectrophotometric detection of formazan, which is produced by cells from 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) in direct correlation to the number of cells present and was measured using a commercially available MTS kit (Promega G3581).

Evaluation of chest CT scans

Chest CT scans were evaluated by a visual scoring method as previously described.[5] Briefly, the extent of groung glass opacification (GGO) of the lung parenchyma, a measure of surfactant accumulation in aPAP,[6-8] was evaluated remotely by two blinded, independent radiologists who utilized dedicated software (Medidata) for the evaluation and recorded results directly into an electronic case report form. The area of lung parenchyma affected by GGO was measured using a zonal sampling of CT scan images including three transverse plane images representing regions of the upper lung (just above the aortic arch), middle lung (at the main carina), and lower lung (at the bifurcation of the lingular and lower lobe bronchi). The area of GGO, which correspondes to the area of involvement according to the following scoring system: 0 = no GGO, 1 = less than 5% GGO, 2 = 5-24% GGO, 3 = 25-49% GGO, 4 = 50-74% GGO, 5 = 75% or more

GGO. Both reviewers assigned a zonal HRCT GGO score for each region and the final GGO score was calculated as the average of all three regions. After completing the GGO score assessment, the reviewers determined if the extent of GGO at week 24 had changed compared to baseline. If reviewers did not agree on the extent of GGO when comparing images, an adjudicator reviewed the change assessment and made a determination without knowing the results of either Reader 1 or Reader 2. The radiologists also noted any adverse findings.

Arterial Blood Gas Measurement

Specific written instructions for blood gas analysis were provided to each Site. Briefly, patients were placed in a supine position for 10 minutes prior to specimen collection. The analysis at Baseline was planned to be conducted with the patient breathing room air when possible. However, if the patient could not tolerate temporary discontinuation of supplemental oxygen during blood gas sampling at Baseline, blood gas sampling at subsequent visits should be conducted using the same oxygen flow rate as was used at Baseline. Blood specimens for arterial blood gas analysis at Baseline (Visit 2) and week 24 (Visit 8) were required to be obtained by arterial puncture. Use of capillary blood specimens was permitted for other times (Visit 1, 3, 5, and follow-up visits) for sites experienced in use of capillary specimens for blood gas analysis. The sample was analysed in accordance with local laboratory practices.

The following variables were measured: PaO_2 and $PaCO_2$. The A-aDO_2 was calculated based on PaO_2, $PaCO_2$ using a following equation:

$$Aa\ Gradient = \left(F_iO_2(P_{atm} - P_{H_2O}) - \frac{P_aCO_2}{0.8} \right) - P_aO_2$$

where Aa Gradient is the alveolar – arterial difference in oxygen concentration (A-aDO_2), F_iO_2 is fraction of inspired oxygen, P_{atm} is ambient atmospheric pressure, P_{H2O} is saturated vapour pressure of water, $PaCO_2$ is arterial partial pressure of carbon dioxide, and PaO_2 is arterial partial pressure of oxygen. The atmospheric pressure was determined at the time each test was performed. The FiO_2 was assumed to be 0.21, and the P_{H20} was assumed to be 47 mm Hg (the value at normal body temperature).

Pulmonary Function Testing

Standardized pulmonary function testing was performed according to guidelines established by the American Thoracic Society/European Respiratory Society (ATS/ERS) Task Force[9-12] by laboratory personnel with documented training in lung function testing. Atmospheric temperature and pressure were measured in parallel with pulmonary function testing. The lung function variables measured included forced expiratory volume in 1 second (FEV_1), forced vital capacity (FVC), vital capacity (VC) and the diffusion capacity for carbon monoxide (DLCO); all variables were expressed as a percentage of the predicted value – i.e., FEV_1%, FVC%, VC%, and DLCO%. DLCO was adjusted for hemoglobin concentration. During the double-blind treatment period, FEV_1%, FVC%, and VC% were measured using a FlowScreen® spirometer (eResearchTechnology GmbH, Estenfeld, Germany) provided by the sponsor results were read centrally. DLCO was performed using local equipment and results were read centrally. During the open-label treatment period, spirometry and DLCO were measured using local equipment.

Six-Minute Walk Test

The 6MWT was performed according to guidelines established by the ATS/ERS Task Force[13] by technicians with documented training and experience of performing the 6MWT. Whenever possible, the 6MWT was conducted with the patient breathing room air. If the patient required oxygen supplementation at rest, an oxygen titration procedure was followed as part of the 6MWT at the Screening Visit in order to

determine the amount of oxygen supplementation required for the patient to complete the test. The same flow of oxygen should then be used at the patient´s subsequent tests in the trial, if possible.

Disease Severity Score

The aPAP disease severity score (DSS) developed by Inoue[14] is based on the presence of PAP-related symptoms and the degree of reduction in arterial oxygen concentration (PaO_2) determined with the individual breathing room air in the supine position. The DSS uses a 5-point scale to measure aPAP disease severity as follows: 1 = no symptoms and $PaO_2 \geq 70$ mmHg, 2 = symptomatic and $PaO_2 \geq 70$ mmHg, 3 = 60 mmHg $\leq PaO_2 < 70$ mmHg, 4 = 50 mmHg $\leq PaO_2 < 60$ mmHg, 5 = $PaO_2 < 50$ mmHg.

Functional Health Status

Patients were asked to complete the St Georges Respiratory Questionnaire (SGRQ)[15] (translated into the local language for each country) using 4-week recall. The SGRQ was originally designed to measure impact on health status in patients with obstructive airway disease but most of the questions were considered relevant for the evaluation of patients with aPAP.

Reporting of Adverse Events

All enrolled patients were closely monitored for the occurrence of adverse events (AEs) from Screening (Visit 1) through completion of 48 weeks (Visit 13) or 72 weeks (Visit 15). Investigators were asked to collect AEs using a non-leading question such as “have you experienced any new health problems or worsening of existing conditions” as well as reporting events directly observed or spontaneously volunteered by patients. Clearly related signs, symptoms and abnormal diagnostic procedure results were grouped together and reported as a single diagnosis or syndrome whenever possible. All AEs including but not limited to events reported by the patient, or reported in answer to an open-ended question by the investigator or member of this team, which fall into any of the above definitions were required to be recorded as an AE in the eCRF and should include the following information: 1) brief description of the event (diagnosis), 2) start date (and time, if relevant), 3) stop date (and time, if relevant) (or resolution), 4) severity, 5) action taken regarding trial drug, 6) opinion on causality, 7) seriousness, and 8) outcome.

Reporting of Serious Adverse Events

The investigator was responsible for ensuring that all serious adverse events (SAE) were reported to the sponsor immediately, but in any event no later than 24 hours of any site staff becoming aware of the event. SAEs should be reported from the time the informed consent had been signed up to 30 days after the last visit.

Statistical Analysis

Determination of Sample Size

The sample size calculation was revised twice during the trial. Based on the initial sample size calculation for the primary endpoint, 42 randomized subjects (14 in each treatment group) were required, but a total of 51 subjects were planned to be randomized. Following input from the US Food and Drug Administration and a change in key secondary endpoints and SAP, the sample size was increased to 90 randomized subjects (30 in each treatment group). Furthermore, a pre-planned fully blinded sample size re-estimation procedure was conducted in January 2018 for the key secondary endpoints to assess the standard deviations of the 6MWT-Distance and SGRQ, as well as the overall WLL event rate. From each ANCOVA model, the respective residual variance for 6MWD (and SGRQ) were used to calculate sample size based on the effect sizes of 50 m and 10 points, respectively. Similarly, the sample size for comparing WLL rates of 5% (active) versus 20% (placebo) at 24 weeks was evaluated, based on the observed (pooled) WLL rate.

All calculations were made based on a 2-sided 5% significance levels, and 80% and 90% power. Based on this, it was decided to increase the sample size to 135 randomized subjects (45 in each treatment group).

Testing Hierarchy

Confirmatory analyses followed assessment of the primary efficacy endpoint (change in A-aDO_2 from baseline to week 24) and included the three key secondary efficacy endpoints (change in SGRQ Total score and 6MWT-distance from baseline to week 24, and time from randomization to the first use of WLL). The Type I error rate for the two active administration arms and four confirmatory endpoints was controlled by a combination of hierarchical testing and Hochberg adjustment in stepwise fashion as described below and illustrated (Fig. S1 in the Supplemental Appendix).

Step 1) The primary endpoint A-aDO_2 was tested comparing the continuous molgramostim group with the placebo group at the 0.05 alpha level.

Step 2) The three key secondary endpoints were tested comparing the continuous molgramostim group with the placebo group using a truncated Hochberg adjustment procedure with truncation factor 0.75. This would mean ordered alpha levels of 0.0417, 0.0229, 0.0167.

Step 3) The primary endpoint A-aDO_2 was tested comparing the intermittent molgramostim group with the placebo group at an alpha level depending on number of significant findings in the previous step.

Step 4) The 3 key secondary endpoints were tested comparing the intermittent molgramostim group with the placebo group using a Hochberg adjustment procedure. The alpha levels in this step would depend on the alpha level in the previous step. The testing procedure would stop at the first step where the analysis at that step reached no statistical significance. Analyses of all other endpoints were considered supportive and not part of the hierarchical testing procedure and were not used to infer significance.

Statistical Methodology

For the analyses of all endpoints measured as change from baseline to 24 weeks, the value of the parameter at Baseline (Visit 2) evaluated was included in an analysis of covariance (ANCOVA) model. Adjustment for region (Japan versus all other countries) and stratification for the occurrence of WLL within two months before randomization was also included in all analysis models. Statistical analysis involved all three treatment arms in the same model even though the confirmatory testing procedure tested the two active treatment groups separately. Analysis of the time from randomization to first use (or need for) WLL was done using a log-rank test, adjusting for the randomization WLL strata, and a Cox regression as sensitivity analysis. The frequency of WLL was analyzed using a negative binominal model including the same covariates as the primary analysis model and using log of study participation time as offset. SGRQ Total Score responder analysis was performed using logistic regression models including the same covariates as were used for endpoints measured as change from baseline to 24 weeks.

For continuous variables, empirical mean values and standard deviations are presented in-text and in tables. Estimated differences of continuous molgramostim versus placebo are presented as least square means with 95% confidence intervals (95%CI) and results with two-sided tests. Categorical outcomes are presented as ratios with 95%CIs.

As specified in the SAP, a treatment policy/de-facto estimand approach was applied for the primary analysis that included all observed data irrespective of subject adherence to the randomized treatment. Patients who withdrew from their randomized treatment before week 24 were followed up through to week 24, and their week 24 data were considered as primary, irrespective of treatment discontinuation, treatment interruptions, or use of rescue medication (WLL or other).

The primary endpoint was also analyzed using a revised full analysis set (FAS) identical to the pre-specified FAS except that the results for A-aDO_2 (and PaO_2) for 4 patients (1 in each molgramostim group and 2 in the placebo group) were invalid, treated as missing data, and replaced by multiple imputation. In these 4 patients, the blood gas measurement was done while they were breathing oxygen via nasal canula, thus, the true A-aDO_2 could not be calculated because the FiO_2 was unknown (see below, Section 4, Supplemental Results and Discussion).

Subjects who withdrew from the study early and who did not have week 24 data available had measurements imputed using multiple imputation. This was done by first ensuring a monotone missing data pattern by imputation of intermediate missing data (baseline, Week 4 and Week 12) by a Markov chain Monte Carlo (MCMC) approach, where imputations are done by treatment group Japan vs other countries, and WLL (stratification) if possible (strata may be dropped depending on the number of observations per combination). The imputation was done generating 10 data sets. When monotone missingness had been accomplished, then the remaining missing data, up to and including Week 24, was imputed using a monotone regression approach using treatment and WLL (stratification) as class variables. Each replication of the data, including the imputed missing values, was analysed using the ANCOVA used in the primary analysis, and estimates was pooled using proc MIANALYZE SAS®: this approach is valid under assumption that data is missing according to 'missing at random'. The imputation of missing data was not described in the SAP, where the primary analysis used a no imputation of missing data approach with the MI procedure applied as a sensitivity analysis for primary and key secondary endpoints. In the revised-FAS analysis of the primary endpoint, missing data from patients breathing supplemental oxygen during arterial blood gas measurement were treated as missing data and replaced by multiple imputation as described above.

Evaluation of residuals from statistical modeling did not reveal important deviations from the normality assumption for the data used for analysis of the primary endpoint. Importantly, the revised FAS was used for only for arterial blood gas-based endpoints (A-aDO_2, PaO_2) and not for the analysis of any other endpoints.

Sensitivity analyses were performed for the primary endpoint (change in A-aDO_2) and key secondary endpoints (SGRQ Total Score and 6MWD). Analyses assumed that missing data were missing at random for analysis by the method of mixed model for repeated measures [MMRM]. A nonrandom missing data assumption was used for worst case imputation and tipping point analysis (only done for primary endpoint) and for the per protocol and completer analyses. The MMRM model used the same covariates as for the primary model; an unstructured covariance matrix was used to model the within-subject errors. This methodology did not require or use imputation of missing data; instead, missing data were modelled based on the subjects' available data and on other subjects' data over time. Multiple imputation was used for patients who had WLL performed as rescue treatment (MI following rescue). For the patients who received WLL during the study, all A-aDO_2 measurements following WLL were excluded and re-imputed using multiple imputation. For the worst-case imputation, patients in either of the active arms had the value of the 10^{th} percentile from their dose arm imputed (i.e. approximately the 3^{rd} worst case) while patients in the placebo arm had the median for the placebo arm imputed. The completer analysis included only data for subjects who completed 24 weeks of blinded intervention and who did not discontinue treatment. Tipping point analysis was performed using primary analysis model. The primary analysis was also repeated for the per-protocol population (ie. patients without major protocol deviation evaluated prior to un-blinding to be included in the per-protocol population (not shown).

SECTION 5. Supplemental Results and Discussion

GM-CSF autoantibody titers and serum neutralization capacity

All enrolled patients had confirmed positive samples and titers reported (Table. 2 in the manuscript). One patient was randomized in error and was not confirmed to have autoimmune PAP as shown by the absence of GM-CSF autoantibodies[16,17] at Baseline (Visit 2). All serum specimens positive for GM-CSF autoantibodies were confirmed to have abnormally increased GM-CSF neutralizing capacity (not shown).

Primary Endpoint Analysis

Identification of Arterial Blood Gas Results for Four Patients as Medical and Statistical Outliers

After unblinding the results of the IMPALA trial, one patient was noted to have a large, negative value (-42 mm Hg) for A-aDO_2 that was non-physiological and suggested that oxygen concentration was higher in arterial blood than alveoli, which is opposite of the oxygen concentration gradient known to exist between these two compartments. Based on this observation, a thorough re-examination of the data revealed that four patients – who required clinically indicated oxygen therapy during the study – had undergone arterial blood gas specimen collection while breathing supplemental oxygen by nasal canula (Table S5 in the Supplemental Appendix). Performance of arterial blood gas analysis in subjects breathing supplemental oxygen via nasal canula precludes calculation of the true A-aDO_2 based on the following considerations. First, in patients breathing supplemental oxygen by nasal canula, the FiO_2 varies by up to 8% depending on the respiratory rate and degree of mouth breathing; and second, the A-aDO_2 varies by up to 7.6 mm Hg for each 1% change in FiO_2.[18] Further, graphic analysis of PaO_2 versus A-aDO_2 for all enrolled patients, i.e., the FAS population (541 measuremts of each variable among 138 patients) revealed the results for the 4 patients who underwent blood gas analysis while breating supplemental oxygen were widely separated from the 134 patients who underwent testing while breathing room air (Fig S3B in the Supplemental Appendix). Therefore, from a medical perspective, the A-aDO_2 results for these four patients could not be considered accurate.

Quantile-quantile analysis of the A-aDO_2 data in the FAS population identified the results for these 4 patients as outliers (Fig. S3C-E in the Supplemental Appendix). Further, robust nonlinear regression analysis of the A-aDO_2 in the FAS population also identified the results for these 4 patients as significant statistical outliers (Fig. S3F in the Supplemental Appendix). Therefore, from a statistical perspective, the A-aDO_2 results for these four patients were outliers.

In summary, from both a medical and statistical perspective the A-aDO_2 results for these four patients were unreliable and, consequently, were treated as 'missing data' for analysis of the primary endpoint. This approach is consistent with the US Food and Drug Administration guidelines for clinical trials, which state: "clear identification of a particular value as an outlier is most convincing when justified medically as well as statistically, and the medical context will then often define the appropriate action".[19] Because PaO_2 is also influenced by the use of supplemental oxygen, analysis of this endpoint was done using the R-FAS (Table S3 in the Supplemental Appendix). Analysis of data for all other endpoints was performed using the FAS according to the prespecified hierarchical testing procedure (Fig. S1 in the Supplemental Appendix). Although not prespecified in the protocol or statistical analysis plan, use of the revised FAS (only for analysis of the primary endpoint) was justified by the strong medical, scientific and statistical rational.

Further Evaluation of Efficacy of Continuous Molgramostim During the Double-Blind Treatment Period

Sensitivity analyses were performed for the primary (Fig. S4) and key secondary endpoints (SGRQ Total Score and 6MWD) (not shown). The results of these sensitivity analyses, done only for the primary endpoint using the R-FAS, were similar to results of the primary analysis and confirmed the main study findings (Fig. S4 in the Supplemental Appendix).

SGRQ Total Score responder analysis showed that there were a higher number of 4-, 8-, and 12-point responders[20] in the continuous molgramostim group compared to placebo (Fig. S5). The Hazard ratio for the time from randomization until the first use of WLL was in favor of continuous molgramostim but was not different from placebo (Fig. S7) as was the rate ratio for frequency of WLL (Fig. S7). Finally, blood hematocrit and hemoglobin levels, which are increased in patients with aPAP as pathophysiologic response to chronic lung disease) declined at a faster rate in patients receiving continuous molgramostim than placebo and leveled out during the open-label extension treatment period (Fig. S9). Together, these results support the conclusion that molgramostim has therapeutic efficacy in patients with aPAP.

Efficacy of Intermittent Molgramostim During the Double-Blind Treatment Period

Effects on Pulmonary Gas Transfer

The mean A-aDO_2 change in patients receiving intermittent molgramostim (-10.6±15.7) or placebo (-7.4±11.1 mmHg) was not significantly different when analyzed using the R-FAS (Fig. S10A and Table S3 in the Supplemental Appendix).

The mean change in DLCO percent of predicted in patients receiving intermittent molgramostim (7.7±11.4 than placebo (3.9±10.9) was not different (Fig. S10B, Table S3 in the Supplemental Appendix).

Effects on Health Status and Functional Exercise Capacity

The mean change in SGRQ total score was greater in patients receiving intermittent molgramostim (-12.0±15.1) than placebo (-4.7±12.8) (Fig. S10C, Table S3 in the Supplemental Appendix). The improvement was associated with greater scores in the activity component and impact component but not the symptoms component (Fig. S10C, Table S3 in the Supplemental Appendix). Responder analysis was different at the 4-point and 12-point but not the 8-point threshold (Fig S5B).

The mean change in distance covered during a six-minute walk test in patients receiving intermittent molgramostim (11.3±81.9) or placebo (6.0±110.5) was not different (Table S3 in the Supplemental Appendix).

Effects on Use of WLL

The time from baseline to the first use (or frequency) of WLL therapy in patients receiving intermittent molgramostim or placebo during the 24-week blinded treatment period was not different (Fig. S7A Table S3 in the Supplemental Appendix). Furthermore, compared to the rate of WLL use before randomization, in the double-blind period, the rate of WLL use was numerically similar in patients receiving placebo and lower in patients receiving intermittent molgramostim (Fig. S7B in the Supplemental Appendix).

Effects on Radiographic and Biochemical Measures of Pulmonary Alveolar Proteinosis Pathology

The reduction in the alveolar surfactant burden from baseline to 24 weeks, which is reflected by the change in GGO score on the chest CT, in patients receiving intermittent molgramostim (-2.1±2.7) or placebo (-1.1±2.5) were not different (ETD -0.8 [95%CI -2.1 to 0.4) (Fig. S10D and Table S3 in the Supplemental Appendix).

Serum levels of some biomarkers of aPAP (LDH, KL-6, CEA, Cfra21-1, and SP-D) but not others (SP-A, SP-B, SP-C) were improved in patients receiving intermittent molgramostim compared to placebo (Fig. S8 in the Supplemental Appendix).

Section 7. Supplemental Figures

Section 6. Supplemental Figures

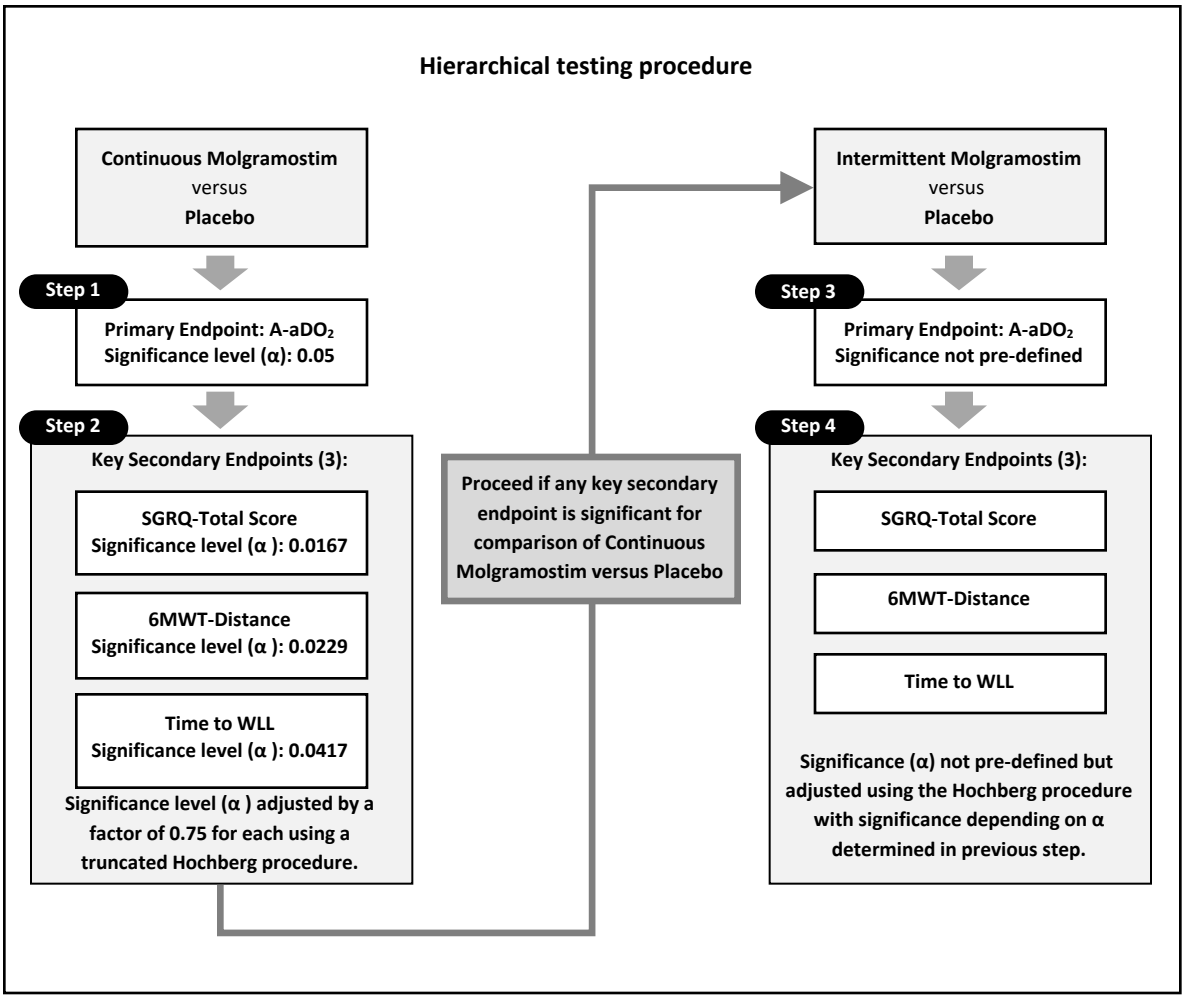

Figure S1A. Prespecified Hierarchical Testing Procedure for Analysis of the IMPALA Trial Results.
The results of the IMPALA trial were evaluated using a stepwise hierarchical testing algorythm with a Truncated Hochberg adjustment procedure to control the Type I error among comparisons of the two active treatment arms with the primary endpoint and three confirmatory key secondary endpoints. See text for details.

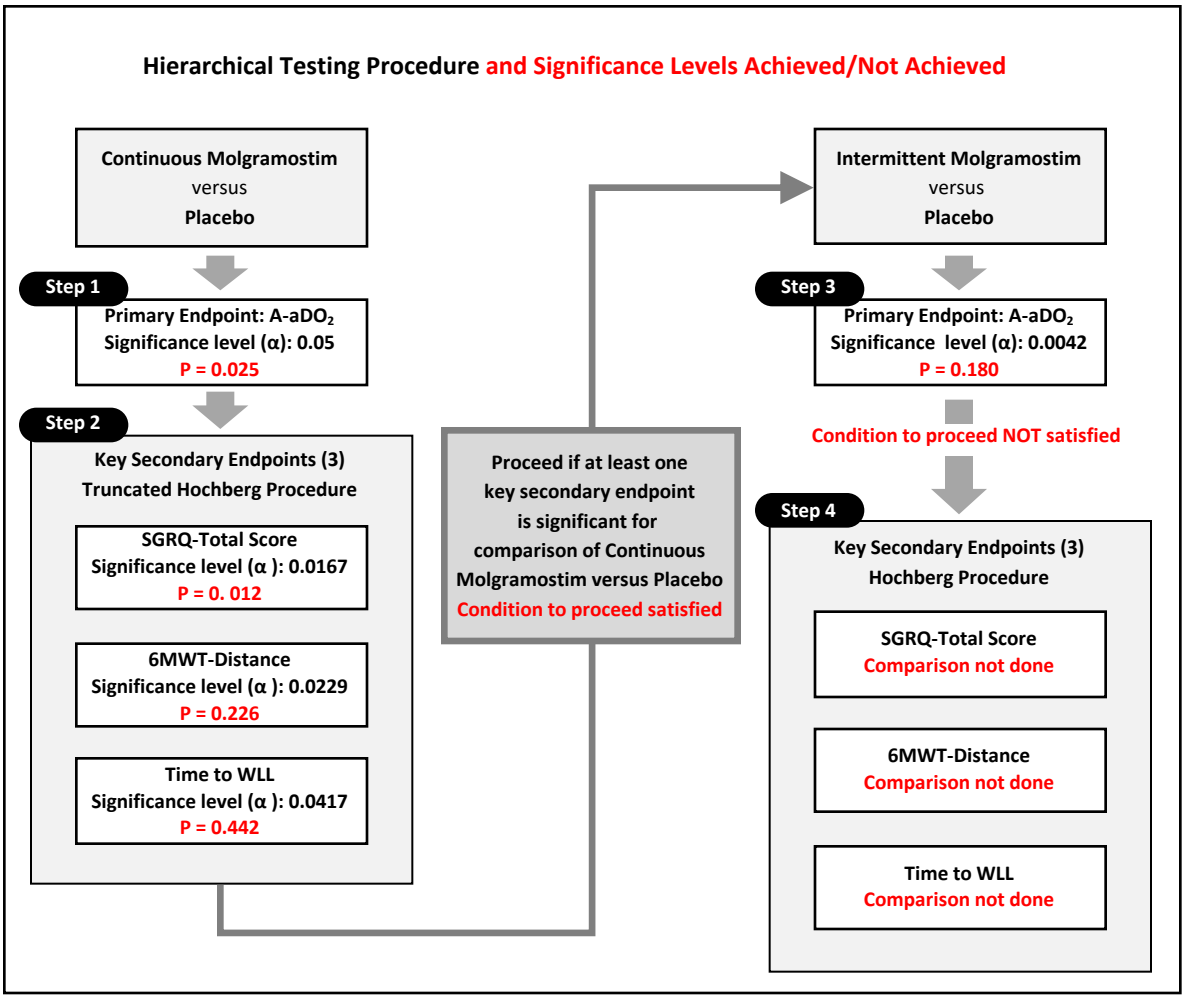

Figure S1B. Results of Significance Testing Using the Heirarchical Testing and Hochberg Testing Procedure.
The IMPALA trial results were compared in stepwise fashion as indicated and the significance values at each step are indicated. In step 2, since one of the null hypotheses was rejected among the key secondary endpoints for comparison of continuous molgramostim versus placebo groups, the hierarchical testing proceeded with comparison of the intermittent molgramostim versus placebo groups for the primary endpoint based on a significance level depending on the number of successes in previous step, as applicable to the truncation factor used. Since the null hypothesis was not rejected for the comparison of primary endpoint in intermittent molgramostim versus placebo, comparison of the key secondary endpoints for intermittent molgramostim vs placebo was not done. See text for further details.

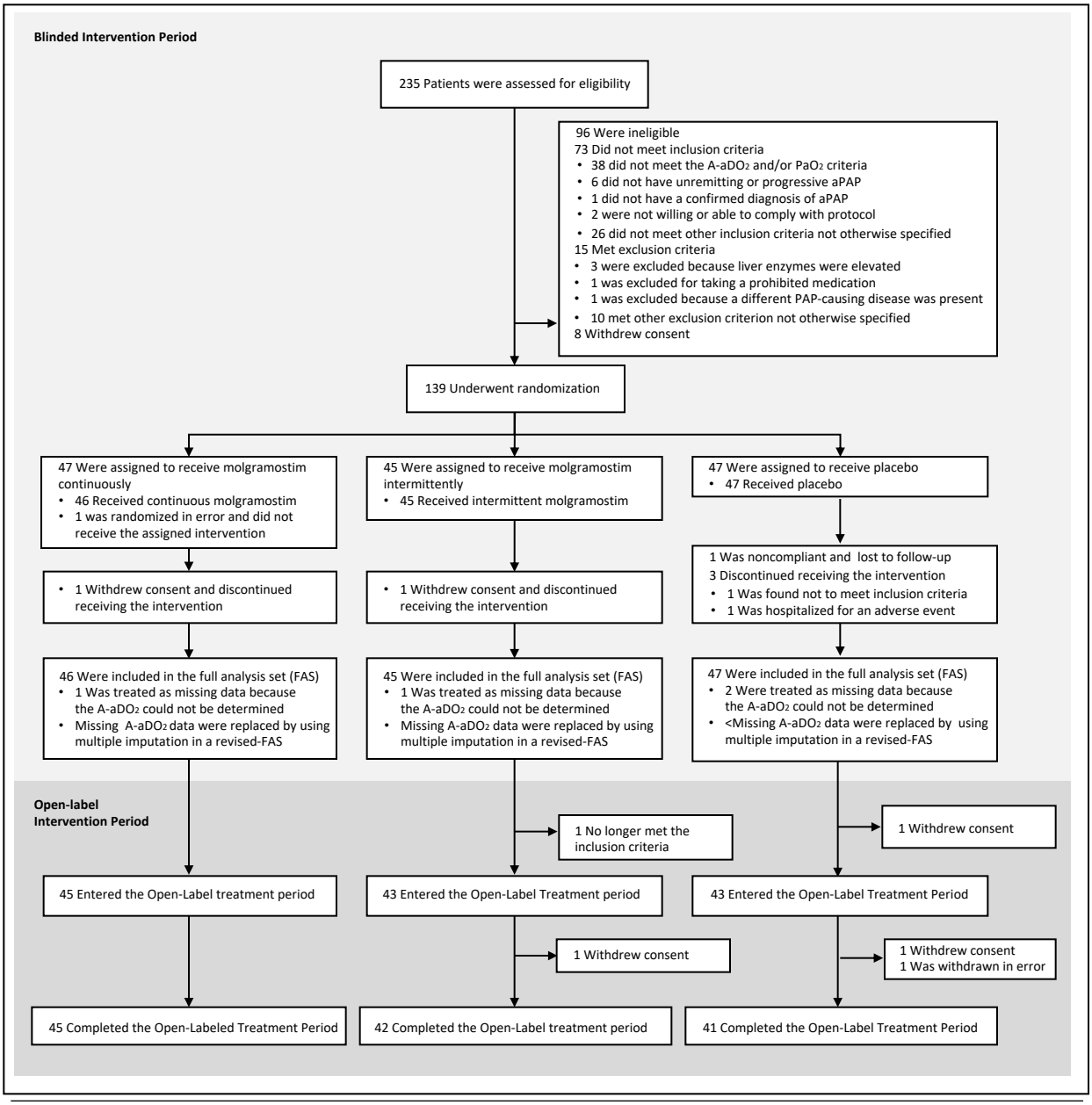

Figure S2. Patient Screening, randomization, and Follow-up.

During the 24-week randomized, double-blind, placebo-controlled intervention period, all patients remained unaware of their trial-group assignments. Of the 235 screening events that occurred (included rescreening of approximately 24 patients), 138 patients with mild – severe aPAP were randomized and received either continuous molgramostim, intermittent molgramostim, or placebo, 133 completed the double-blind period of the trial, and 131 completed the open-label extension treatment period. Assessments were performed every 4 weeks during the double-blind treatment period and every 12 weeks during the open-label extension treatment period. All 138 patients were included in the full analysis set (FAS), and 134 patients were included in the revised FAS (R-FAS) that was used for evaluation of the primary end point.

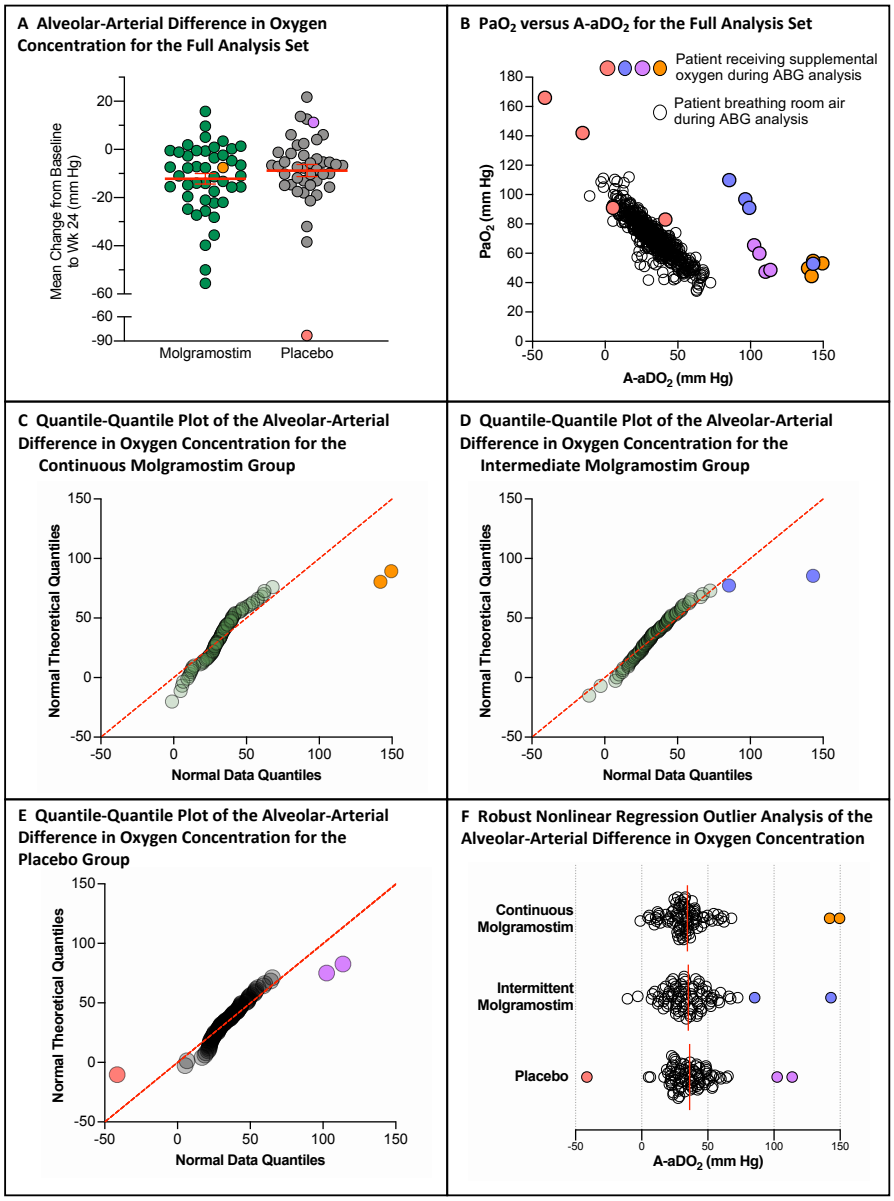

Figure S3. Change in A-aDO$_2$ from Baseline to Week 24 using the Full Analysis Set and Outlier Analysis of the Results of Arterial Blood Gas Measurements.

Patients breathing room air during specimen collection for arterial blood gas analysis are indicated by green circles (continuous molgramostim; Panels A, C, D) or grey circles (placebo Panels A, E) or open circles (either molgramostim or placebo; Panels B, F) while those breathing supplemental oxygen via nasal canula at various rates are indicated by orange circles (continuous molgramostim, n=1; Panels A, B, C, F), Purple or salmon circles (Panel A, B, E, F), or blue circles (Intermittent molgramostim (n=1; Panels B, D, F). Panel A shows the mean (±SEM) change from baseline to week 24 in A-aDO$_2$ in the continuous molgramostim group (46 patients) compared to the placebo group (47 patients) during the randomized double-blind intervention period analyzed using the FAS. Panel B shows the relationship between PaO$_2$ and A-aDO$_2$ results for all arterial blood gas measurements (n=541) in all enrolled study patients (n=138) including patients breathing room air (open circles) or supplemental oxygen via nasal canula (colored circles) during the collection of blood specimens for analysis. Each circle represents one measurement in one patient. Each color represents a different patient. Panels C-E show quantile-quantile plots of the distribution of A-aDO$_2$ results (Normal Data Quantiles) compared to a normal distribution (Normal Theoretical Quantiles) for results at baseline (Visit 2) and week 24 (Visit 8) for patients in the continuous molgramostim group (n=46, Panel C), intermittent molgramostim group (n=45, Panel D), and placebo group (n=47, Panel E). Panel F shows robust nonlinear regression outlier analysis of the A-aDO$_2$ results for all patients in the indicated groups.

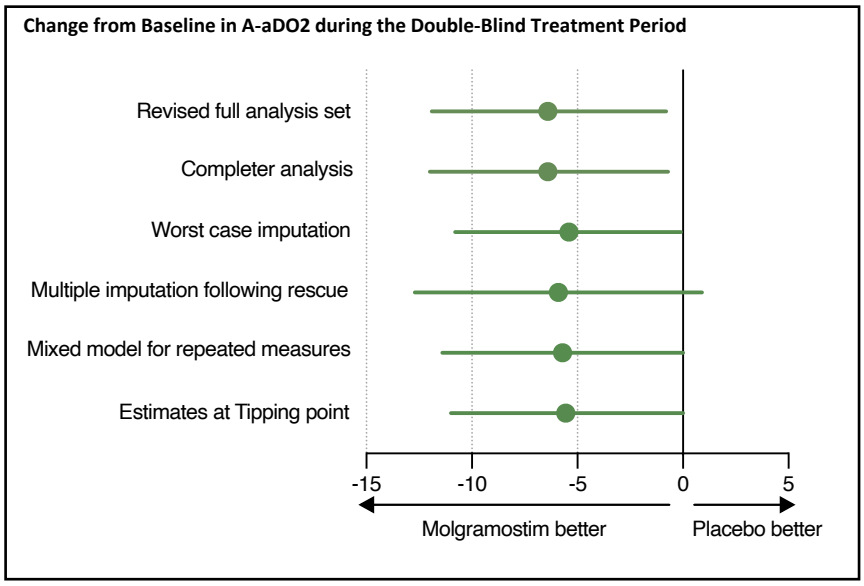

Figure S4. Sensitivity Analysis for the Primary Endpoint During the Double-Blind Period.
Least squares mean (± 95% confidence intervals) values are shown for each sensitivity analysis for the comparison of the change from baseline at 24 weeks in A-aDO_2 between the continuous molgramostim group (46 patients) and placebo group (47 patients) performed using the R-FAS.

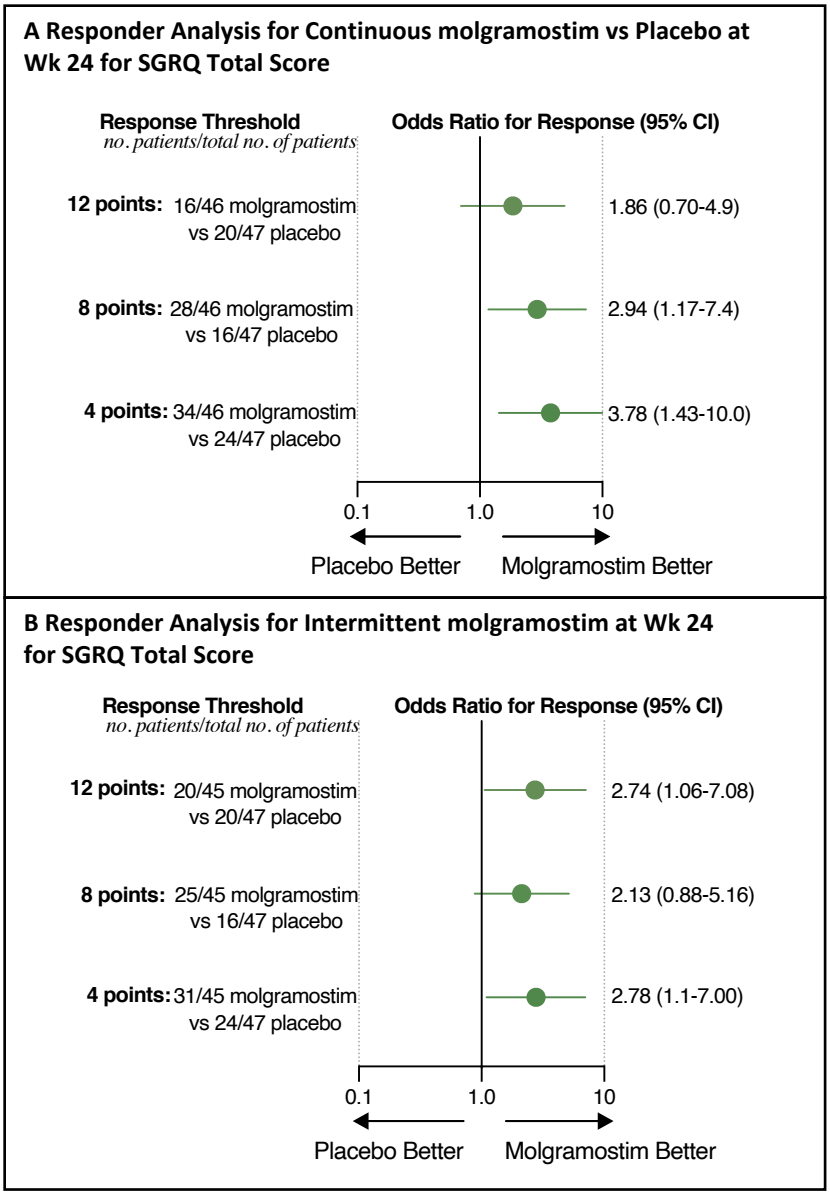

Figure S5. Responder analysis for the comparison of SGRQ total scores in patients receiving molgramostim versus placebo during the double-blind period.

Panel A shows odds ratio and 95% confidence interval (95% CI) for a response for the patients receiving continous molgramostim (n=46) or placebo (n=47) at the indicated response threshold levels (4, 8, or 12 points) of the SGRQ total score. Panel B shows similar analysis of the results for patients recieving intermittent molgramostim (n=45) compared to placebo (n=47).

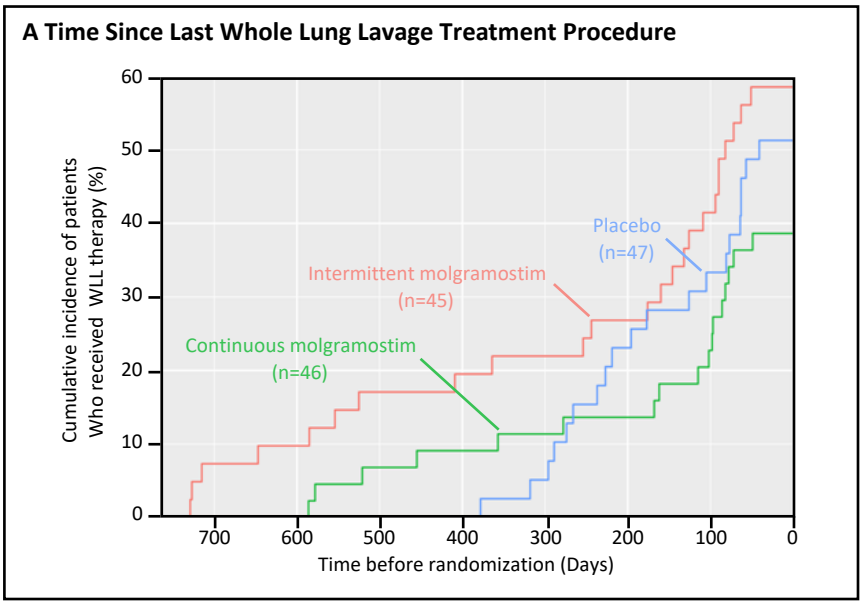

Figure S6. Comparison of the Prior use of Whole Lung Lavage Therapy Among the Three Intervention Groups. Panel A shows a Kaplan-Meier plot of the percentage of the 138 patients in the three randomized intervention groups who underwent whole lung lavage prior to enrollment in this study.

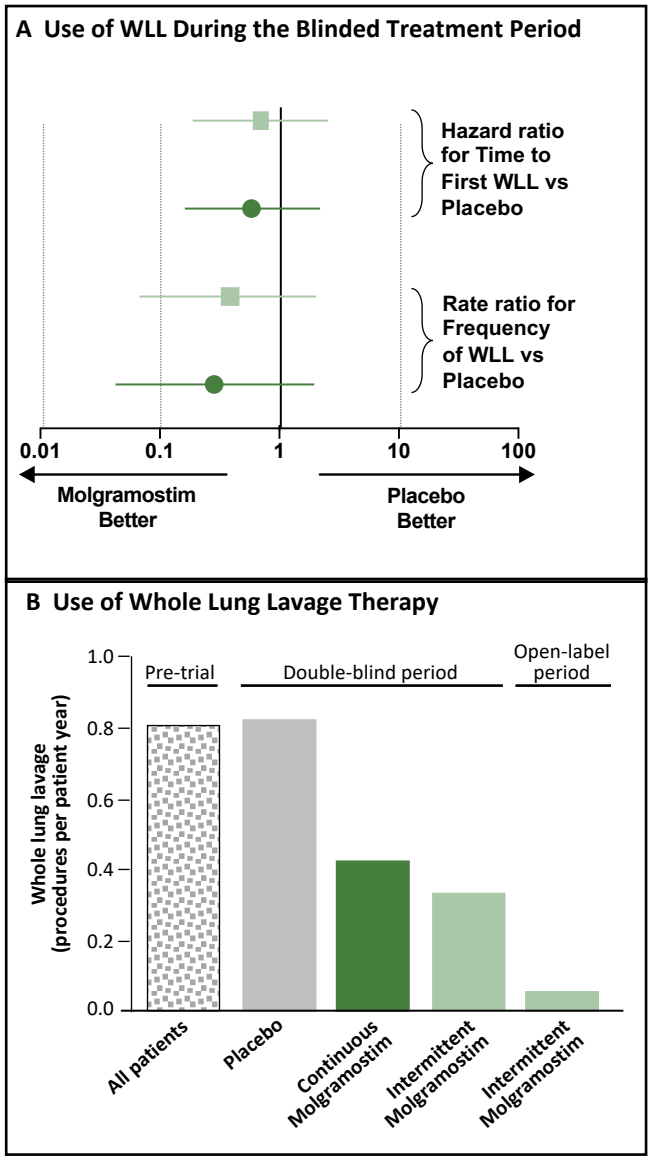

Figure S7. Analysis of the use of Whole Lung Lavage Therapy Before and During the IMPALA Trial.
Panel A shows the hazard ratio for the time to first use of WLL and rate ratio for the freuency of WLL use in patients receiving continuous molgramostim or intermittent molgramostim compared to placebo during the double-blind period. Panel B shows the rate of WLL use, expressed in number of single-lung WLL procedures per patient year before randomization in the IMPALA trial (Pre-trial), in the patients in receiving continous molgramostim (n=46), intermittent molgramostim (n=45), or placebo (n=47) during the randomized treatment period (double-blind period), and in the patients (n=131) enrolled in the open-label treatment extension period (Open-label period). Data analyzed using the FAS.

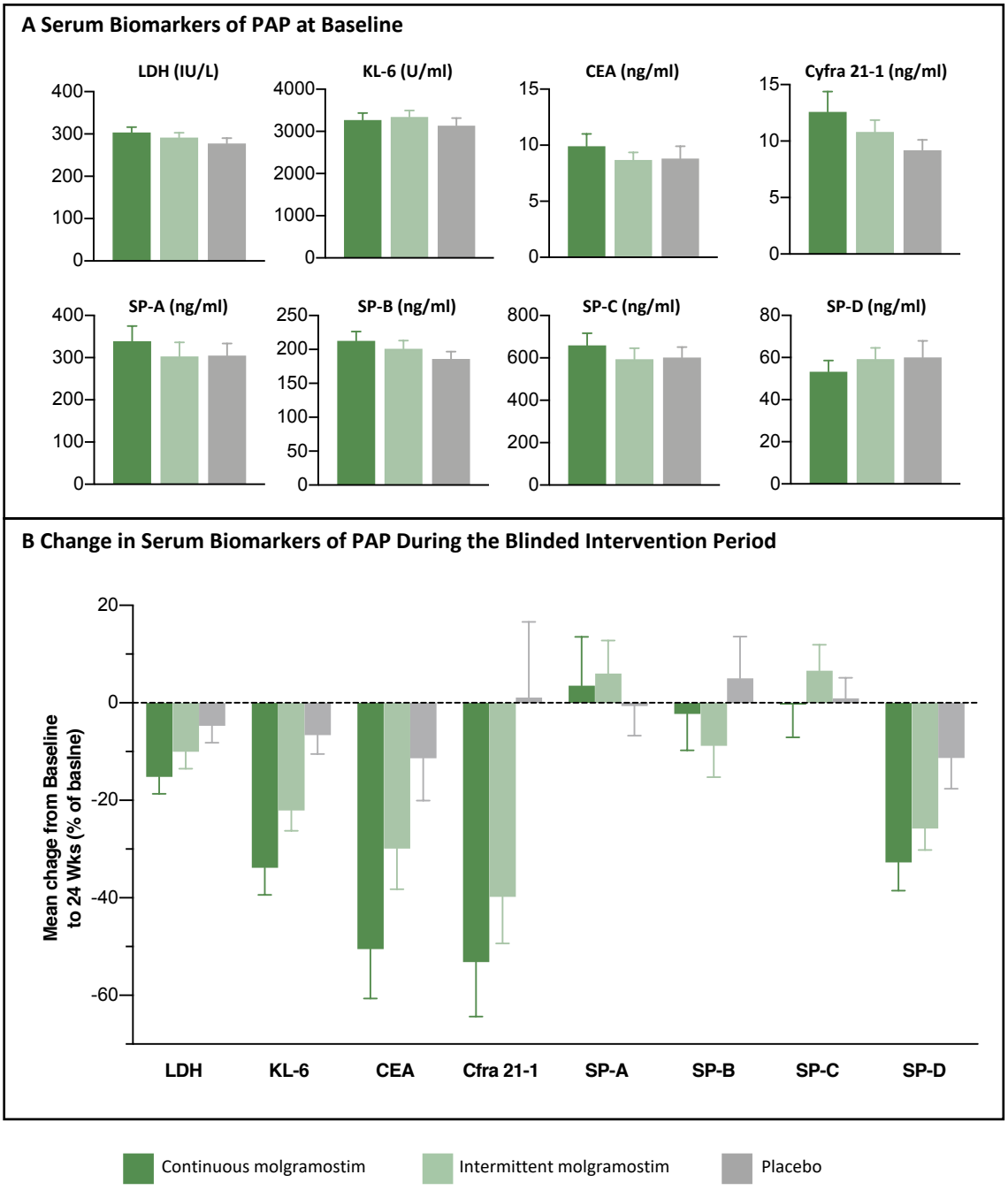

Figure S8. Change in Serum Biomarker Levels During the Double-Blind Intervention Period.
Shown are mean (±SEM) values at Baseline (A) and mean (±SEM) percentage change from baseline at 24 weeks for serum biomarkers of aPAP including lactate dehydrogenase (LDH), Krebs von den Lungen-6 (KL-6), carcinoembryonic antigen (CEA), cytokeratin fragment 19 (Cyfra 21-1), surfactant protein A (SP-A), surfactant protein B (SP-B), surfactant protein C (SP-C), and surfactant protein D (SP-D) in the patients receiving continuous molgramostim, intermittent molgramostim, or placebo (indicated). Analysis was performed on the FAS of observed values without imputation of missing data.

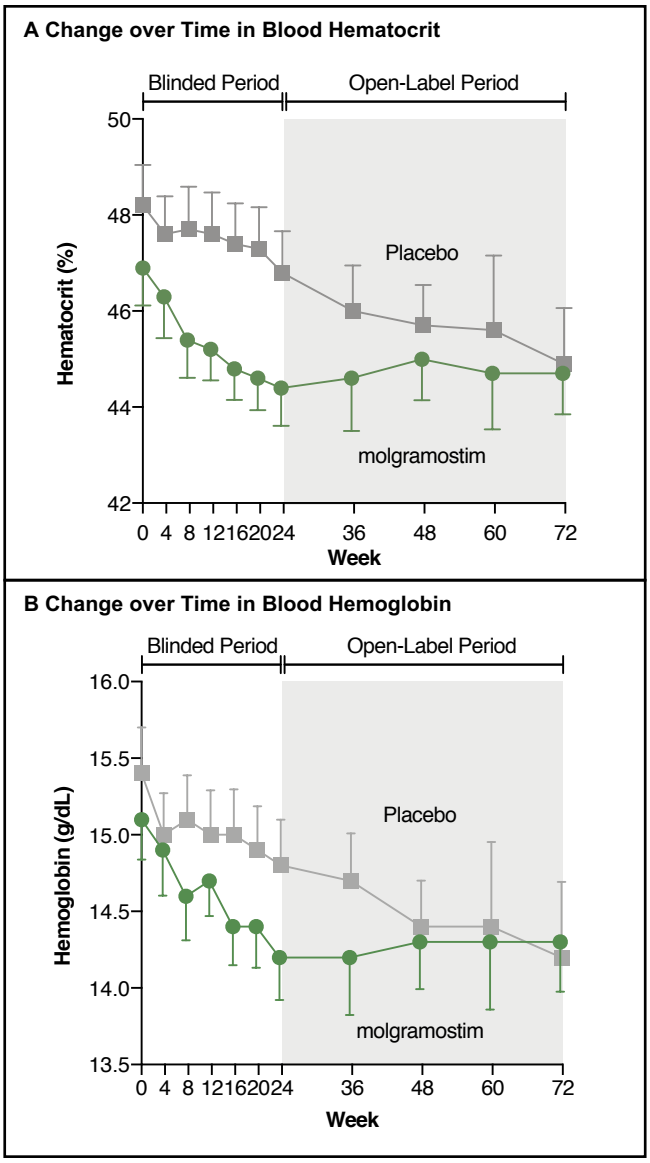

Figure S9. Change Over Time in Blood Hematocrit and Hemoglobin Levels in Patients Receiving Continuous Molgramostim or Placebo During the IMPALA Trial.

Panel A shows the mean (±SEM) hematocrit in patients receiving continuous molgramostim (n=46) or placebo (n=47) during randomized treatment period (Blinded period) or open-label extension treatment period (n=131) (Open-label period). Data are shown for the FAS of observed values without imputation of missing data.

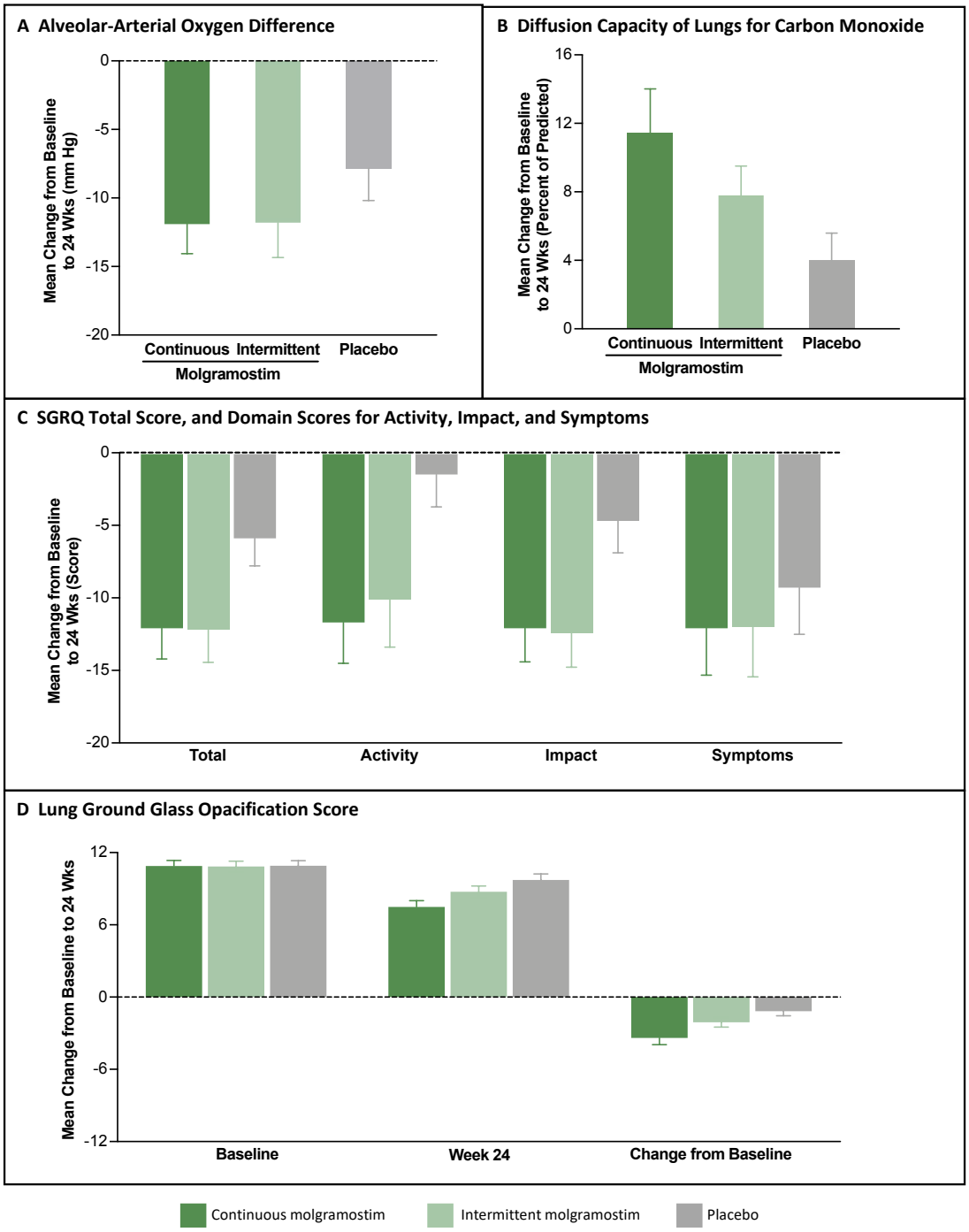

Figure S10. Changes in A-aDO_2, DLCO%, SGRQ Scores, and GGO Scores in Patients Receiving Molgramostim or Placebo During the Double-Blind Treatment Period.

Results are mean (±SEM) changes for each endpoint in patients receiving intermittent molgramostim (green bars) vs placebo (gray bars); for comparison results for patients receiving continuous molgramostim are shown (open bars). Panel A shows the results for changes in the alveolar-arterial difference in oxygen concentration (A-aDO_2). Panel B shows the results for changes in DLCO%. Panel C shows the results for SGRQ total score, and activity, impact, and symptom components. Panel D shows the results for chest CT GGO scores at baseline (Visit 2), week 24 (Visit 8), and the change from baseline to week 24. Panel A was analyzed using the R-FAS and Panels B-D with the FAS with imputation of missing data except for the SGRQ component scores, which include only observed values without immutation.

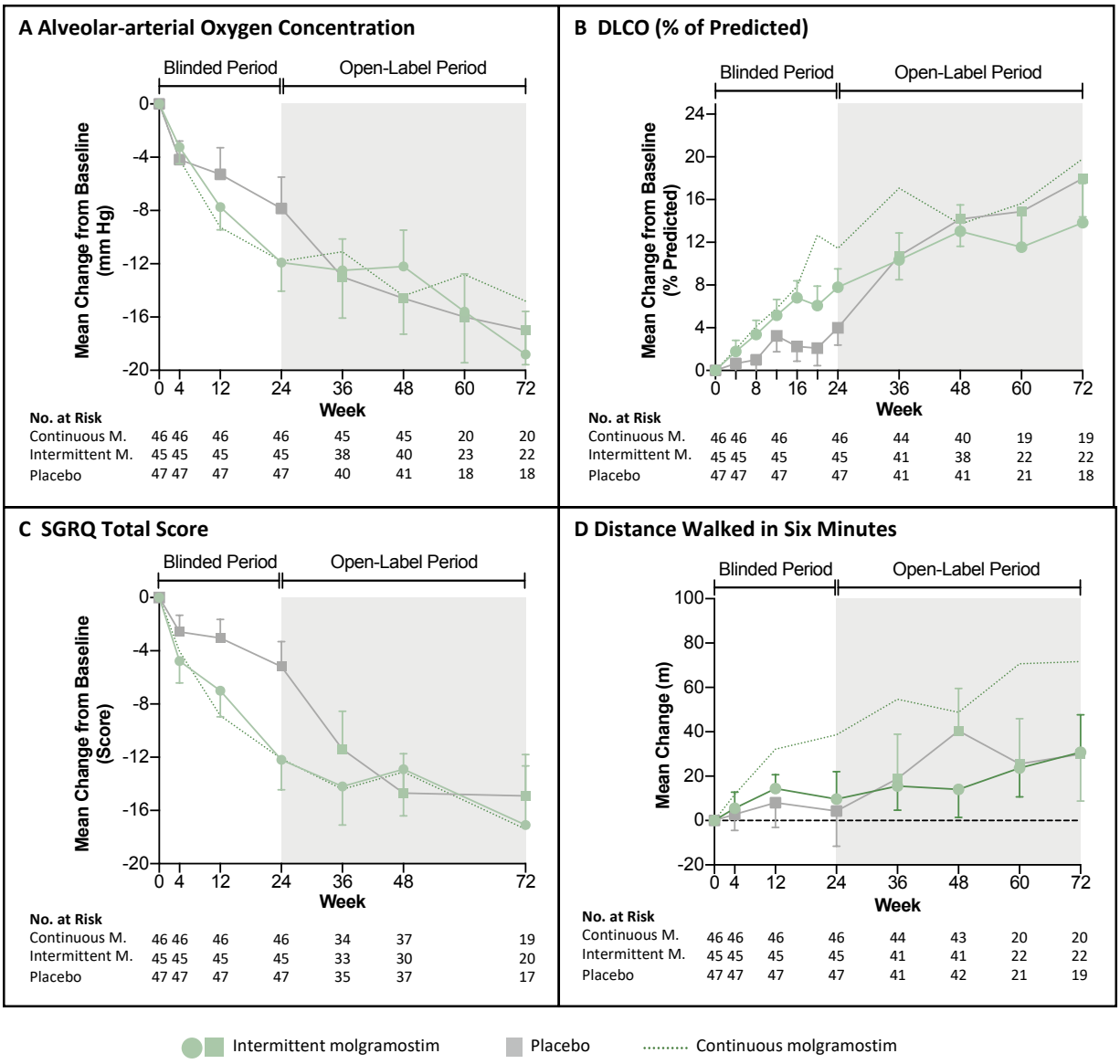

Figure S11. Changes from Baseline Over Time in A-aDO_2, DLco, SGRQ total score, and 6MWT-distance in Patients Receiving Intermittent Molgramostim or Placebo During the Double-Blind and Open-Label Treatment Periods. Panels A-D show the mean (±SEM) changes from baseline to week 24 (Double-blind Period, white regions) and from week 24 to week 72 (Open-label Period, grey regions) for A-aDO_2 (Panel A), DLCO% (Panel B), SGRQ total score (Panel C), and the distance covered in a six-minute walk test analyzed using the R-FAS (Panel A) or FAS (Panels B-D). Panel A was analyzed using the R-FAS and Panels B-D with the FAS with imputation of missing data.

Section 7. Supplemental Tables

Table S1. Schedule of Events and Procedures for the IMPALA Trial. *

Visit Type	Screening	Baseline	Double-Blinded Treatment Period						Open-Label Treatment Period				
Visit number	1	2	3	4	5	6	7	8	9	10	11	12 [a]	13 [a]
Study week	-2	0	4	8	12	16	20	24	28 [b]	36	48	60	70
Study day (± visit window)	-14±7	1±7	28±7	56±7	84±7	112±7	140±7	168±7	196±3	204±14	216±14	228±14	238±14
Written informed consent [c]	+												
Eligibility criteria	+	+ [d]											
Demographics	+												
Medical history	+												
Prior/concurrent medications [e]		+	+	+	+	+	+	+					
Treatment of aPAP									+	+	+	+	+
Pregnancy test/contraceptive check [f]	+	+	+	+	+	+	+	+		+	+	+	+
Physical Exam	+		+		+			+			+		+
Vital signs	+		+		+			+			+		+
Arterial blood gas measurement [g]	+	+	+		+			+		+	+	+	+
Lung function tests [h]	+	+	+	+	+	+	+	+		+	+	+	+
6-min walk test [i]	+	+	+		+			+		+	+	+	+
Laboratory safety tests [j]	+	+	+	+	+	+	+	+		+	+	+	+
Electrocardiogram		+			+			+			+		+
GM-CSF [k]		+	+										
Anti-Drug antibody test [c] [l]	+	+	+		+			+		+	+		
Biomarker tests [m]		+	+		+			+					
Chest computed tomography scan [n]		+						+					
Disease severity score	+							+			+		+
Dyspnea score [o]	+	+	+		+			+		+	+	+	+
Cough scores		+	+		+			+		+	+		+
Quality of life score		+	+		+			+		+	+		+
Study drug administration training	+[p]												
Randomization		+											
Administer study drug in clinic		+	+		+			+		+	+		
Collect and record adverse events		+	+	+	+	+	+	+	+	+	+	+	+
Dispense/return diary cards [q]		+	+	+	+	+	+	+		+	+ [r]	+	+
Dispense study drug		+	+	+	+	+	+	+	+	+	+ [r]	+	
Treatment compliance			+	+	+	+	+	+		+	+	+	+

* Trial events are indicated by +. A-aDO_2 denotes alveolar-arterial difference in oxygen concentration, aPAP autoimmune pulmonary alveolar proteinosis, PaO_2 partial pressure of arterial oxygen.

a Visit occured only for patients who provided written informed consent before the approval of protocol amendment number 11.

b Visit conducted by telephone interview.

c Written informed consent was obtained before performing any study procedures except for phlebotomy and GM-CSF autoantibody testing prior to the screening visit to establish a diagnosis of aPAP. Patients who have not had this test previously and at sites that conducted a chest CT before screening.

d Eligibility criteria for trial participation was reconfirmed.

e Prior refers to any medication or therapy administered before the first administration of any blinded treatment intervention in this trial. Concomitant refers to any medication or therapy administered after the first administration of blinded treatment intervention in this trial and before completion of the study.

f A serum pregnancy test was performed during all on-site study visits (Visit 1-8, 10-13) for participants receiving blinded study medication. A urine pregnancy test was performed immediately before the first administration of blinded study medication (Visit 2) to confirm the participant is not pregnant and meets this inclusion criterion. Pregnancy testing (urine or serum) was performed monthly while administering inhaled molgramostim during the open-label period (Visits 10-13); pregnancy testing was not mandatory for participants who did not receive molgramostim during the open label-period.

g Performed on blood specimens obtained by arterial puncture at baseline (Visit 2) and completion of the blinded treatment period (Visit 8); testing at other times (Visits 1, 3-7, 10-13) was performed on arterial or capillary blood.

h Includes measurement of vital capacity (VC), diffusion capacity of the lungs for carbon monoxide (DLco), forced expiratory volume in 1 second (FEV_1), and forced vital capacity (FVC), all of which were expressed as the percentage of the respective predicted values.

i The requirement for administration of supplemental oxygen therapy during conduct of the 6-minute walk test (as defined in the protocol) was determined at Screening or Baseline (Visits 1 or 2). For participants requiring such therapy, the administration of oxygen via nasal canula was maintained at the same flow rate during conduct of this test at all subsequent study visits.

j Includes routine clinical hematology and chemistry tests.

k Serum specimens for measurement of GM-CSF concentration were obtained immediately prior to and 2 hours after the first administration of blinded study drug at Baseline (Visit 2) and one month later (Visit 3).

l Serum specimens for measurement of anti-drug antibodies were obtained immediately prior to administration of blinded study drug.

m Includes serum concentrations of lactate dehydrogenase (LDH), mucin-like glycoprotein (KL-6), carcinoembryonic antigen (CEA), cytokeratin 19 fragment (Cyfra 21-1), surfactant protein (SP) A-D (SP-A, SP-B, SP-C, SP-C).

n Allowed to be obtained up to four weeks before Baseline and up to 1 week before or after completion of the double-blinded treatment period (Visit 8). If obtained before Baseline as part of this study, written informed consent was collected before Screening (Visit 1)

o Assessed using the Borg dyspnea scale immediately before and immediately after performing a six-minute walk test.

p All participants received training in the use and maintenance of the study drug delivery device before administration; the time and location of the training was determined at Screening (Visit 1).

q A new (blank) diary card was provided to each participant at each study visit beginning at Baseline (Visit 2) and collected at the following visit during the blinded intervention period (Visits 3-8) and during the open-label period (Visits 10-13). Participants were asked to record on their diary card information related to adverse events and to answer questions about potential suspected unexpected potential lung or systemic effects.

r Applicable only to enrolled participants who provided written informed consent prior to the approval of protocol amendment number 11.

Table S2. Clinical Characteristics of the Patients at Baseline.*

Characteristic	Continuous Molgramostim Group (N=46)	Intermittent Molgramostim Group (N=45)	Placebo Group (N=47)
Age – yr	54.0±13.3	49.2±14.1	46.1±14.8
Female gender – no. (%)	18 (39.1)	19 (42.2)	22 (46.8)
Tobacco use – no. (%)			
Never smoker	13 (28.3)	16 (35.6)	16 (34.0)
Ex-smoker	27 (58.7)	20 (44.4)	20 (42.6)
Current smoker	6 (13.0)	9 (20.0)	11 (23.4)
Clinical and laboratory features			
Time since diagnosis of PAP – months	39.8±58.1	40.0±45.9	32.0±31.5
Serum GM-CSF autoantibody titer †	68,561±113,710	45,364±54,474	66,756±110,907
Hematocrit – %	46.9±5.2	48.8±5.7	48.2±5.7
Hemoglobin – g/dL	15.1±1.73	15.5±1.78	15.3±1.86
Pulmonary gas exchange			
A-aDO_2			
No. of patients with data (R-FAS §)	45	44	45
Mean value – mm Hg ‡	38.1±10.8	38.6±12.9	38.8±11.2
PaO_2			
No. of patients with data (R-FAS §)	45	44	45
Mean value – mm Hg	65.5±9.9	66.0±11.4	67.1±13.2
Percent of predicted DLCO	52.1±18.6	46.1±14.5	49.6±14.3
Pulmonary airflow and lung volumes			
Percent of predicted FEV_1	89.3±23.7	82.5±22.4	79.3±17.7
Percent of predicted FVC	83.0±21.5	80.4±12.2	80.2±16.7
Percent of predicted VC	78.6±3.2	74.8±19.5	74.1±18.6
Disease severity score (DSS) – no. (%) ¶			
DSS-1	4 (8.7)	5 (11.1)	3 (6.4)
DSS-2	12 (26.1)	14 (31.1)	16 (34.0)
DSS-3	17 (37.0)	13 (28.9)	14 (29.8)
DSS-4	5 (10.9)	9 (20.0)	10 (21.3)
DSS-5	8 (17.4)	3 (6.7)	4 (8.5)
Pulmonary surfactant accumulation			
No. of patients with data	46	43	47
CT GGO score ‖	10.9±3.2	10.8±3.0	10.9±2.8
Functional health status			
Saint Georges Respiratory Questionnaire **			
Total score	47.2±20.4	44.4±21.4	44.1±21.7
Activity domain score	60.7±21.5	57.2±25.4	55.6±22.6
Impact domain score	38.9±23.6	35.3±22.3	35.8±23.5
Symptom domain score	46.3±22.7	47.1±25.4	47.1±24.3
Exercise capacity			
No. of patients with data	45	45	47
Distance walked on 6-min walk test (m)	412±144	447±117	447±125
Prior or concomitant therapy of PAP			
Supplemental oxygen – no. (%)	15 (32.6)	12 (26.7)	12 (25.5)
WLL therapy††			
Any prior WLL – no. of patients (%)	23 (50.0)	31 (68.9)	30 (63.8)
No. of prior WLL procedures	3.8	3.7	2.8
Time since last WLL procedure – mo	24.3±52.6	18.9±24.0	17.7±20.7
GM-CSF therapy			
Any prior GM-CSF – no. of patients (%)	6 (13.0)	7 (15.6)	6 (12.8)
Time since last administration – mo	35.4±35.9	37.8±26.4	18.3±22.6

* Plus-minus values are means ± SD. Results represent data for N=46, 45, or 47 patients in the Continuous Molgramostim, Intermittent Molgramostim, and Placebo groups, respectively, unless indicated otherwise; data include observed values only without any imputation of missing or invalid data. A-aDO_2 denotes alveolar-arterial oxygen difference, CT computed tomography, DLCO diffusing capacity for carbon monoxide, DSS [PAP] disease severity score, FEV_1 forced expiratory volume in 1 second, FVC forced vital capacity, GGO ground glass opacification score, GM-CSF granulocyte/macrophage-colony stimulating factor, R-FAS revised full analysis set, VC vital capacity, WLL whole lung lavage.

† Determined as the actual to nominal concentration of GM-CSF autoantibodies from serial 2-fold dilutions of serum measured by quantitative electro-chemiluminescence assay.

‡ The alveolar-arterial difference in oxygen concentration is calculated with the use of the following equation: $A\text{-}aDO_2 = FiO_2 \times (PB\text{-}P_{H20}) \times PaCO_2/R) - PaO_2$, where FiO_2 indicates fraction of inspired oxygen, $PaCO_2$ partial pressure of arterial carbon dioxide, PB barometric pressure measured by validated barometers, P_{H20} partial pressure of water vapor in inspired air (assumed to be 47 mm Hg), and R the respiratory quotient (assumed to be 0.8).

§ In the revised full analysis set (R-FAS), $A\text{-}aDO_2$ results for four patients were invalid and treated as missing data. See text for details.

¶ The PAP disease severity score ranges from 1 -5; DSS-1 = no symptoms and $PaO_2 \geq 70$ mm Hg, DSS-2 = symptomatic and $PaO_2 \geq 70$ mm Hg, DSS-3 = $PaO_2 < 70$ and ≥ 60 mm Hg, DSS-4 = $PaO_2 < 60$ and ≥ 50 mm Hg, DSS-5 = $PaO_2 < 50$ mm Hg.

‖ The GGO scores range from 0 to 15, with higher scores indicating a higher proportion of the area of the chest CT scan images corresponding to lung parenchyma affected by ground glass opacification, an indication of the abnormal accumulation of surfactant sediment in patients with aPAP.

** Scores on the SGRQ range from 0 to 100, with higher scores indicating more severe effects on a patient's functional health status.

†† The lifetime utilization of WLL therapy of PAP before the baseline visit, including any prior use, the number of single lung WLL treatments performed, and the time elapsed since the most recent WLL treatment.

‡‡ The lifetime utilization of GM-CSF as therapy of PAP before the baseline visit.

Table S3. Effects of Molgramostim on Primary and Selected Secondary Outcome Variables After 24 Weeks of Treatment.*

	Value at 24 Weeks			Change from Baseline			Estimated Difference	
Variable	Continuous Molgramostim (N = 46)	Intermittent Molgramostim (N = 45)	Placebo (N = 47)	Continuous Molgramostim (N = 46)	Intermittent Molgramostim (N = 45)	Placebo (N = 47)	Continuous Molgramostim vs. Placebo (95% CI) †	Intermittent Molgramostim vs. Placebo (95% CI) †
Pulmonary gas exchange								
$A\text{-}aDO_2$								
No. of patients with data (R-FAS) ‡	44	42	41	44	42	41		
$A\text{-}aDO_2$ (mm Hg)	26.1±13.7	27.9±15.9	31.8±12.5	-12.2±14.7	-10.6±15.7	-7.4±11.1	-6.4 (-11.9 to -0.8)	-3.4 (-9.0 to 2.2)
PaO_2								
No. of patients with data (R-FAS) ‡	44	42	41	44	42	41		
PaO_2 (mm Hg)	76.5±14.4	76.2±14.9	73.1±12.9	11.2±14.0	10.4±16.0	6.4±11.7	7.4 (-0.3 to 15.8)	4.4 (-3.1 to 12.5)
DLCO								
No. of patients with data	44	43	42	43	43	42		
Percentage of predicted DLCO	63.3±22.9	54.7±12.3	54.4±15.8	11.6±17.3	7.7±11.4	3.9±10.9	7.9 (2.2 to 13.6)	2.9 (-2.8 to 8.7)
Pulmonary airflow and lung volumes								
No. of patients with data	45	42	43	45	42	43		
Percent of predicted FEV_1	91.4±23.9	83.4±19.9	80.5±18.8	1.9±11.5	0.5±8.9	0.6±12.5	2.3 (-2.2 to 6.9)	0.4 (-4.2 to 4.9)
Percent of predicted FVC	86.8±21.8	82.9±17.0	81.9±18.2	3.5±11.0	2.1±8.7	1.8±10.0	1.7 (-2.4 to 5.8)	0.2 (-3.9 to 4.3)
Percent of predicted VC	83.1±22.1	78.3±18.9	75.3±20.6	4.1±11.1	3.3±9.2	1.7±12.3	2.7 (-1.9 to 7.2)	1.8 (-2.7 to 6.4)
Radiological evaluation of the lungs								
Number of patients	44	44	43	44	43	43		
GGO score §	7.5±3.7	8.8±3.2	9.7±3.4	-3.4±3.8	-2.1±2.7	-1.1±2.5	-2.4 (-3.7 to – 1.2)	-0.8 (-2.1 to 0.4)
Functional health status								
Saint Georges Respiratory Questionnaire¶								
Number of patients	45	44	43	45	44	43		
Total score	35.4±21.3	31.7±18.4	37.8±24.2	-12.3±14.3	-12.0±15.1	-4.7±12.8	-7.6 (-13.4 to -1.8)	-7.0 (-12.7 to -1.3)
Activity domain score	49.4±26.8	46.2±24.6	52.5±27.6	-11.7±18.9	-10.1±28.8	-1.5±14.6	-10.4 (-18.3 to -2.4)	-7.9 (-15.7 to -0.1)
Impact domain score	27.4±21.8	22.2±16.7	29.5±24.2	-12.1±15.4	-12.4±15.7	-4.7±14.4	-6.9 (-12.7 to -1.0)	-7.5 (-13.4 to -1.7)
Symptom score	34.9±22.5	34.9±25.9	36.4±29.5	-12.1±21.6	-12.0±22.9	-9.3±21.1	-3.1 (-11.9 to 5.7)	-2.6 (-11.4 to 6.1)
Exercise capacity								
No. of patients with data	44	43	43	43	43	43		
Distance walked in six minutes, m	448±136	468±111	459±144	39.6±96	11.3±82	6.0±111	20.6 (-19.8 to 61.0)	5.6 (-34.1 to 45.2)
Molgramostim/GM-CSF antibody titer								
No. of patients with data	45	43	43	43	43	43		
Serum titer	63499±102978	77258±34641	64250±95040	-2989±119314	32061±124787	7287±57041	-	-

* Plus-minus values are empirical means ± SD; data include observed values only without any imputation of missing or invalid data. CI denotes confidence interval.

† Between-group differences for change from baseline are expressed as least square means and 95% confidence intervals with use of an analysis of covariates model (all treatment groups included in the same model) with treatment, WLL within 2 months prior to Baseline (stratification) and geographic region (Japan vs other countries) as factors and Baseline values as covariates.

‡ In the revised full analysis set (R-FAS), $A\text{-}aDO_2$ results for four patients were invalid and treated as missing data. See text for details.

§ The GGO scores range from 0 to 15, with higher scores indicating a higher proportion of the area of the chest CT scan images corresponding to lung parenchyma affected by ground glass opacification, an indication of the abnormal accumulation of surfactant sediment in patients with aPAP.

¶ Scores on the SGRQ range from 0 to 100, with higher scores indicating more severe effects on a patient's functional health status.

Table S4. Effect on the Primary End Point of Excluding the Four Patients Who Received Supplemental Oxygen Therapy During Arterial Blood Gas Measurement *

	Value at Baseline (Week 0)		Value at Week 24		Change from Baseline		Estimated Difference	P-value
Data Set – Imputation Use A-aDO_2†	Continuous Molgramostim	Placebo	Continuous Molgramostim	Placebo	Continuous Molgramostim	Placebo	Continuous Molgramostim vs. Placebo (95% CI)	
FAS – no imputation ‡								
No. of patients with data	46	47	45	43	45	43		
A-aDO_2, mm Hg	40.5± 19.6	40.2±14.3	28.6±22.0	32.0±20.9	-12.1± 14.6	-8.8±16.1	-4.6 (-11.1 to 2.0)	0.169
R-FAS – no imputation §								
No. of patients with data	45	45	44	41	44	41		
A-aDO_2, mm Hg	38.1±10.8	38.8±11.2	26.1±13.7	31.8±12.5	-12.2±14.7	-7.4±11.1	-6.4 (-11.9 to -0.8)	0.025
FAS – imputation of missing data¶								
No. of patients with data	46	47	46	47	46	47		
A-aDO_2, mm Hg	40.5± 19.6	40.2±14.3	28.6± 21.8	32.4±20.7	-11.9±14.6	–7.8±16.1	-5.2 (-11.6 to 1.3)	0.118
R-FAS – imputation of missing and invalid data‖								
No. of patients with data	46	47	46	47	46	47		
A-aDO_2, mm Hg	38.2±10.9	38.6±11.2	26.4±13.7	31.6±12.7	-11.9±14.9	-7.0±11.4	-6.2 (-11.7 to -0.8)	0.025

* These four individuals excluded from the analysis of the primary endpoint (but not other analyses) included 1 patient in the continuous molgramostim group, 2 patients in the placebo group, and 1 patient in the intermittent molgramostim group (See Table S5). Data are mean (±SD) A-aDO_2 in mm Hg.

† The primary endpoint, the alveolar-arterial difference in oxygen concentration (A-aDO_2), was calculated with the use of the following equation: A-aDO_2 = FiO_2 x (PB-P_{H20}) – $PaCO_2$/R) – PaO_2, where FiO_2 indicates the fraction of inspired oxygen, $PaCO_2$ the partial pressure of arterial carbon dioxide, PB the barometric pressure measured by validated barometers, P_{H20} the partial pressure of water vapor in inspired air (assumed to be 47 mm Hg), and R the respiratory quotient (assumed to be 0.8). A-aDO_2 was pre-specified to be measured with patients at rest and in a supine position while breathing room air for at least 10 minutes – when permitted based on the patients' clinical condition.

‡ Calculated using the measured A-aDO_2 results for all patients with available data (indicated).

§ Calculated using the measured A-aDO_2 results for all patients with available data except for the A-aDO_2 results for four patients that were determined to be invalid because blood gas measurement was done while supplemental oxygen was being administered via nasal canula and, thus, the true A-aDO_2 could not be calculated because the FiO_2 was unkown; results for these four patients were treated as missing data and omitted from the revised FAS without imputation.

¶ Calculated using results for each patient receiving at least one dose of intervention, with replacement of missing values by multiple imputation; invalid data for the four patients breathing supplemental oxygen during arterial blood gas measurement were included and not replaced by multiple imputation.

‖ Calculated using results for each patient receiving at least one dose of intervention, with replacement of missing data by multiple imputation; invalid data for the four patients breathing supplemental oxygen during arterial blood gas measurement were also replaced by multiple imputation.

Table S5. Results of Selected End Points for the Four Patients for whom the Arterial-Alveolar Difference in Oxygen Concentration Results Were Judged to be Invalid.*

		A-aDO_2 (mm Hg)		PaO_2 (mm Hg)		DLCO (% Predicted)		SGRQ Total Score		6MWT-Distance (m)		WLL (No.)
Patient†	Intervention Group‡	Baseline	Week 24	Baseline	Week 24	Baseline	Week 24	Baseline	Week 24	Baseline	Week 24	DB-Period
1	Continuous molgramostim	149.5	142.0	53.2	44.5	23.1	24.4	88.2	85.0	81§¶	102§¶	0
2	Placebo	41.5	-41.5**	83.0	166.0	61.2	71.5	49.4	16.2	532	571	0
3	Placebo	102.5	113.7	65.5	48.7	39.6	NR	88.2	88.4	260¶	165¶	0
4	Intermittent molgramostim	143.1	85.4	52.8	109.9	34.4	66.2	26.4	20.8	350¶	435¶	0

* Shown are individual results for each patient at the indicated time points; data are measured values. A-aDO_2 denotes alveolar-arterial difference in oxygen concentration, PaO_2 arterial blood oxygen concentration, DLCO diffusion capacity of the lungs for carbon monoxide, NR not recorded, SGRQ Saint Georges Respiratory Questionnaire, 6MWT-Distance distance walked during a six minute walk test, WLL whole lung lavage.

† To maintain confidentiality, arbitrary patient numbers were assigned that do not identify the clinical site or the unique identification number assigned to the patient during the trial.

‡ Shown is the treatment group in which each patient participated during the double-blind intervention period of the study.

§ The 6MWT was discontinued early for safety because the SpO_2 decreased significantly during the test.

¶ Supplemental oxygen therapy was administered during the 6MWT in order for the patient to maintain an adequate PaO_2 during the test.

** Large negative values of A-aDO_2 are not physiologically plausible since they require that arterial oxygen concentration to be higher than alveolar oxygen concentration.

Table S6. PAP Disease Severity Score Assigned to Study Patients.*

Observation Time	Disease Severity Score†	All Patients (n=138)	Continuous Molgramostim (n=46)	Intermittent Molgramostim (n=45)	Placebo (n=47)
Screening Visit (Wk 0)					
	1	12 (8.7%)	4 (8.7%)	5 (11.1%)	3 (6.4%)
	2	42 (30.4%)	12 (26.%)	14 (31.1%)	16 (34.0%)
	3	44 (31.9%)	17 (37.0%)	13 (28.9%)	14 (29.8%)
	4	24 (17.4%)	5 (10.9%)	9 (20.0%)	10 (21.3%)
	5	15 (10.9%)	8 (17.4%)	3 (6.7%)	4 (8.5%)
	Missing Data	1 (0.7%)	0 (0%)	1 (2.2%)	0 (0%)
End of Double-blind Period (Wk 24)					
	1	37 (26.8%)	16 (34.8%)	12 (26.7%)	9 (19.1%)
	2	46 (33.3%)	12 (26.1%)	15 (33.3%)	19 (40.4%)
	3	32 (23.2%)	11 (23.9%)	12 (26.7%)	9 (19.1%)
	4	8 (5.8%)	4 (8.7%)	2 (4.4%)	2 (4.4%)
	5	8 (5.8%)	2 (4.3%)	2 (4.4%)	4 (8.5%)
	Missing Data	7 (5.1%)	1 (2.2%)	2 (4.4%)	4 (8.5%)

* Data are numbers (%) of patients assigned the indicated Disease Severity Score (DSS) for all enrolled patients and randomized to the three treatment groups during the blinded intervention period. PAP denotes pulmonary alveolar proteinosis.

† The PAP disease severity score ranges from 1 -5; DSS-1 = no symptoms and PaO_2 ≥70 mm Hg, DSS-2 = symptomatic and PaO_2 ≥70 mm Hg, DSS-3 = PaO_2 < 70 and ≥ 60 mm Hg, DSS-4 = PaO_2 < 60 and ≥50 mm Hg, DSS-5 = PaO_2 < 50 mm Hg.

Table S7. Adverse Events in the Three Study Groups during the Double-Blind Treatment Period*

Category	Adverse Events No. of patients (%) Continuous molgramostim (n=46)	Intermittent molgramostim (n=45)	Placebo (N=47)
Any adverse event†	39 (84.8)	41 (91.1)	41 (87.2)
Most common adverse events‡			
Cough	15 (32.6)	12 (26.7)	11 (23.4)
Chest pain	10 (21.7)	2 (4.4)	1 (2.1)
Headache	6 (13.0)	7 (15.6)	7 (14.9)
Nasopharyngitis	7 (15.2)	10 (22.2)	6 (12.8)
Dyspnea	5 (10.9)	7 (15.6)	4 (8.5)
Productive cough	4 (8.7)	3 (6.7)	3 (6.4)
Pain in extremity	4 (8.7)	0	0
Progression of aPAP§	3 (6.5)	5 (11.1)	8 (17.0)
Weight increase	3 (6.5)	5 (11.1)	0
Nausea	3 (6.5)	1 (2.2)	1 (2.1)
Pyrexia	2 (4.3)	3 (6.7)	3 (6.4)
Back pain	2 (4.3)	3 (6.7)	1 (2.1)
Arthralgia	2 (4.3)	0	4 (8.5)
Chest discomfort	1 (2.2)	3 (6.7)	1 (2.1)
Peripheral edema	1 (2.2)	3 (6.7)	1 (2.1)
Diarrhea	0	6 (13.3)	3 (6.4)
Upper respiratory tract infection	0	3 (6.7)	3 (6.4)
Oxygen saturation decreased	0	0	3 (6.4)
Adverse events possibly or probably related to the intervention	15 (32.6)	11 (24.4)	14 (29.8)
Adverse events leading to discontinuation of the intervention	2 (4.3)	1 (2.2)	1 (2.1)
Total number of adverse events	215	191	192

* Values are the numbers of patients that experienced the indicated adverse events. Values in parentheses are the percentage of patients in the treatment group that experienced the indicated adverse events. A patient with multiple occurrences of a specific adverse event was counted only once in the specific adverse event category. aPAP denotes autoimmune pulmonary alveolar proteinosis.

† Adverse events were coded using the Medical Dictionary of Regulatory Activities, version 21.0.

‡ Adverse events occurring in at least 5% of patients in a treatment group are shown.

§ Progression or worsening of aPAP lung disease was reported as an adverse event.

Table S8. Serious Adverse Events in the Three Study Groups during the Double-Blind Treatment Period*

	Serious Adverse Events No. of patients (%)		
	Continuous molgramostim (n=46)	Intermittent molgramostim (n=45)	Placebo (N=47)
Any serious adverse event†	8 (17.4)	5 (11.1)	8 (17.0)
All serious adverse events			
Progression of aPAP‡	3 (6.5)	3 (6.7)	6 (12.8)
Bacterial pneumonia	1 (2.2)	1 (2.2)	0
Cough	1 (2.2)	0	0
Dyspnea	1 (2.2)	0	0
Laryngeal edema	1 (2.2)	0	0
Respiratory failure	1 (2.2)	0	0
Pneumonia	1 (2.2)	0	0
Respiratory tract infection	1 (2.2)	0	0
Aphasia	1 (2.2)	0	0
Epilepsy	1 (2.2)	0	0
Diverticulitis	0	1 (2.2)	0
Asthma	0	0	1 (2.1)
Lower respiratory tract infection	0	0	1 (2.1)
Gambling disorder	0	0	1 (2.1)
Drug detoxification	0	0	1 (2.1)
Serious adverse events possibly or probably related to intervention	2 (4.3)	0	0
Serious adverse events leading to discontinuation of intervention	1 (2.2)	0	1 (2.1)
Total number of serious adverse events	13	5	16

* Values are the numbers of patients that experienced the indicated adverse events. Values in parentheses are the percentage of patients in the treatment group that experienced the indicated adverse events. A patient with multiple occurrences of a specific adverse event was counted only once in the specific adverse event category. aPAP denotes autoimmune pulmonary alveolar proteinosis.

† Adverse events were coded using the Medical Dictionary of Regulatory Activities, version 21.0.

‡ Progression or worsening of aPAP lung disease was reported as an adverse event.

Table S9. Adverse Events during the Open-Label Treatment Extension Period*

Category	Adverse Events No. of patients (%) Intermittent molgramostim (n=130)
Any adverse event†	87 (66.9)
Most common adverse events‡	
Nasopharyngitis	24 (18.5)
Cough	11 (8.5)
Progression of aPAP§	7 (5.4)
Rash	6 (4.6)
Bronchitis	5 (3.8)
Upper respiratory tract infection	5 (3.8)
Pneumonia	5 (3.8)
Dyspnea	4 (3.1)
Peripheral edema	4 (3.1)
Back pain	4 (3.1)
C-reactive protein increased	4 (3.1)
Adverse events possibly or probably related to the intervention	11 (8.5)
Adverse events leading to discontinuation of the intervention	1 (0.8)
Total number of adverse events	336

* Values are the numbers of patients that experienced the indicated adverse events. Values in parentheses are the percentage of patients in the treatment group that experienced the indicated adverse events. A patient with multiple occurrences of a specific adverse event was counted only once in the specific adverse event category. aPAP denotes autoimmune pulmonary alveolar proteinosis.

† Adverse events were coded using the Medical Dictionary of Regulatory Activities, version 21.0.

‡ Adverse events occurring in at least 3% of patients receiving open-label treatment are shown.

§ Progression or worsening of aPAP lung disease was reported as an adverse event.

Table S10. Serious Adverse Events during the Open-Label Treatment Extension Period*

Category	Adverse Events No. of patients (%) Intermittent molgramostim (n=130)
Any serious adverse event†	16 (12.3)
All serious adverse events	
Progression of aPAP‡	5 (3.8)
Pneumonia	2 (1.5)
Nasopharyngitis	1 (0.8)
Bronchitis	1 (0.8)
Influenza	1 (0.8)
Bacterial pneumonia	1 (0.8)
Urinary tract infection	1 (0.8)
Oral candidiasis	1 (0.8)
Lung infection	1 (0.8)
Esophageal candidiasis	1 (0.8)
Pulmonary tuberculosis	1 (0.8)
Weight decrease	1 (0.8)
Cataract	1 (0.8)
Prostatitis	1 (0.8)
Cholelithiasis	1 (0.8)
Breast cancer	1 (0.8)
Squamous cell carcinoma of the tongue	1 (0.8)
Serious adverse events possibly or probably related to the intervention	0
Serious adverse events leading to discontinuation of the intervention	1 (0.8)
Total number of serious adverse events	23

* Values are the numbers of patients that experienced the indicated adverse events. Values in parentheses are the percentage of patients in the treatment group that experienced the indicated adverse events. A patient with multiple occurrences of a specific adverse event was counted only once in the specific adverse event category.

† Adverse events were coded using the Medical Dictionary of Regulatory Activities, version 21.0.

‡ Progression or worsening of aPAP lung disease was reported as an adverse event.

SECTION 8. References

1. Walter MR, Cook WJ, Ealick SE, Nagabhushan TL, Trotta PP, Bugg CE. Three-dimensional structure of recombinant human granulocyte-macrophage colony-stimulating factor. J Mol Biol 1992;224:1075-85.
2. Marchese RD, Puchalski D, Miller P, et al. Optimization and validation of a multiplex, electrochemiluminescence-based detection assay for the quantitation of immunoglobulin G serotype-specific antipneumococcal antibodies in human serum. Clin Vaccine Immunol 2009;16:387-96.
3. Tanaka N, Watanabe J, Kitamura T, Yamada Y, Kanegasaki S, Nakata K. Lungs of patients with idiopathic pulmonary alveolar proteinosis express a factor which neutralizes granulocyte-macrophage colony stimulating factor. FEBS Lett 1999;442:246-50.
4. Kitamura T, Tanaka N, Watanabe J, et al. Idiopathic pulmonary alveolar proteinosis as an autoimmune disease with neutralizing antibody against granulocyte/macrophage colony-stimulating factor. J Exp Med 1999;190:875-80.
5. Tazawa R, Trapnell BC, Inoue Y, et al. Inhaled granulocyte/macrophage-colony stimulating factor as therapy for pulmonary alveolar proteinosis. American journal of respiratory and critical care medicine 2010;181:1345-54.
6. Robinson TE, Trapnell BC, Goris ML, Quittell LM, Cornfield DN. Quantitative analysis of longitudinal response to aerosolized granulocyte-macrophage colony-stimulating factor in two adolescents with autoimmune pulmonary alveolar proteinosis. Chest 2009;135:842-8.
7. McCarthy C, Lee E, Bridges JP, et al. Statin as a novel pharmacotherapy of pulmonary alveolar proteinosis. Nat Commun 2018;9:3127.
8. McCarthy C, Bartholmai BJ, Woods JC, McCormack FX, Trapnell BC. Automated Parenchymal Pattern Analysis of Treatment Responses in Pulmonary Alveolar Proteinosis. American journal of respiratory and critical care medicine 2019;199:1151-2.
9. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. Eur Respir J 2005;26:319-38.
10. Macintyre N, Crapo RO, Viegi G, et al. Standardisation of the single-breath determination of carbon monoxide uptake in the lung. The European respiratory journal 2005;26:720-35.
11. Miller MR, Crapo R, Hankinson J, et al. General considerations for lung function testing. Eur Respir J 2005;26:153-61.
12. Wanger J, Clausen JL, Coates A, et al. Standardisation of the measurement of lung volumes. Eur Respir J 2005;26:511-22.
13. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002;166:111-7.
14. Inoue Y, Trapnell BC, Tazawa R, et al. Characteristics of a large cohort of patients with autoimmune pulmonary alveolar proteinosis in Japan. American journal of respiratory and critical care medicine 2008;177:752-62.
15. Jones PW, Quirk FH, Baveystock CM. The St George's Respiratory Questionnaire. Respir Med 1991;85 Suppl B:25-31; discussion 3-7.
16. Sakagami T, Beck D, Uchida K, et al. Patient-derived granulocyte/macrophage colony-stimulating factor autoantibodies reproduce pulmonary alveolar proteinosis in nonhuman primates. Am J Respir Crit Care Med 2010;182:49-61.
17. Sakagami T, Uchida K, Suzuki T, et al. Human GM-CSF autoantibodies and reproduction of pulmonary alveolar proteinosis. N Engl J Med 2009;361:2679-81.
18. Wettstein RB, Shelledy DC, Peters JI. Delivered oxygen concentrations using low-flow and high-flow nasal cannulas. Respiratory care 2005;50:604-9.

19. Annonomous. Guidance for Industry: E9, Statistical Principles for Clinical Trials. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and REsearch ICBER), September 1998 (ICH).
20. St. George's Respiratory Questionnaire (SGRQ). American Thoracic Society, 2016. at https://www.thoracic.org/members/assemblies/assemblies/srn/questionaires/sgrq.php.)